

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

THIS PAGE BLANK (USPTO)


```

; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-955-502-19

Query Match      34.5%; Score 48; DB 10; Length 78;
Best Local Similarity 23.6%; Pred. No. 0.0038;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MRRXXXCXGXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db      1 MSRTIFCTFLQREADGQDFQLYPGELGKRIYNEISKEAWAQWQHKTMLINEKKL 55

RESULT 3
US-09-955-502-15
; Sequence 15, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Salmonella enteritidis
US-09-955-502-15

Query Match      34.5%; Score 48; DB 10; Length 88;
Best Local Similarity 23.6%; Pred. No. 0.0042;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MRRXXXCXGXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db      1 MSRTIFCTYLQRDAEGQDFQLYPGELGKRIYNEISKDAWAQWQHKTMLINEKKL 55

RESULT 4
US-09-955-502-33
; Sequence 33, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-955-502-33

Query Match      34.5%; Score 48; DB 10; Length 78;
Best Local Similarity 23.6%; Pred. No. 0.0038;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MRRXXXCXGXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db      1 MSRTIFCTFLQREADGQDFQLYPGELGKRIYNEISKEAWAQWQHKTMLINEKKL 55

RESULT 5
US-09-955-502-11
; Sequence 11, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Escherichia coli K-12 MG1655
US-09-955-502-11

Query Match      34.5%; Score 48; DB 10; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.0043;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MRRXXXCXGXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db      1 MSRTIFCTFLQREAGQDFQLYPGELGKRIYNEISKEAWAQWQHKTMLINEKKL 55

RESULT 6
US-09-955-502-12
; Sequence 12, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Escherichia coli O157:H7EDL933
US-09-955-502-12

Query Match      34.5%; Score 48; DB 10; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.0043;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MRRXXXCXGXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db      1 MSRTIFCTFLQREAGQDFQLYPGELGKRIYNEISKEAWAQWQHKTMLINEKKL 55
```

```

; ORGANISM: Coxiella burnetii
US-09-955-502-33

Query Match      34.5%; Score 48; DB 10; Length 88;
Best Local Similarity 23.6%; Pred. No. 0.0042;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MRRXXXCXGXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db      1 MTRRIICQKLGKEADALNYSYPGELGERIYNHISEQAWQAWLSHQTMLINEYRL 55

RESULT 5
US-09-955-502-11
; Sequence 11, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Escherichia coli K-12 MG1655
US-09-955-502-11

Query Match      34.5%; Score 48; DB 10; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.0043;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MRRXXXCXGXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db      1 MSRTIFCTFLQREAGQDFQLYPGELGKRIYNEISKEAWAQWQHKTMLINEKKL 55

RESULT 6
US-09-955-502-12
; Sequence 12, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Escherichia coli O157:H7EDL933
US-09-955-502-12

Query Match      34.5%; Score 48; DB 10; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.0043;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MRRXXXCXGXXXXXXXXXXPPXXXXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db      1 MSRTIFCTFLQREAGQDFQLYPGELGKRIYNEISKEAWAQWQHKTMLINEKKL 55
```



```
RESULT 7
US-09-955-502-13
; Sequence 13, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Escherichia coli O157:H7
US-09-955-502-13

Query Match      34.5%; Score 48; DB 10; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.0043;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MXRXXXCXXXXXXXXXXXXXXXXXXPPXXXXXGXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MSRTIFCTFLQREAGQDFQLYPGELGKRIYNEISKDAWAQWQHKTMLINEKKL 55

RESULT 8
US-09-955-502-14
; Sequence 14, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Salmonella paratyphi
US-09-955-502-14

Query Match      34.5%; Score 48; DB 10; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.0043;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MXRXXXCXXXXXXXXXXXXXXXXXXPPXXXXXGXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MSRTIFCTFLQREAGQDFQLYPGELGKRIYNEISKDAWAQWQHKTMLINEKKL 55

RESULT 9
US-09-955-502-16
; Sequence 16, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
```

```
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Salmonella dublin
US-09-955-502-16

Query Match      34.5%; Score 48; DB 10; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.0043;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MXRXXXCXXXXXXXXXXXXXXXXXXPPXXXXXGXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MSRTIFCTYLQDAEQDFQLYPGELGKRIYNEISKDAWAQWQHKTMLINEKKL 55

RESULT 10
US-09-955-502-17
; Sequence 17, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Salmonella typhi CT18
US-09-955-502-17

Query Match      34.5%; Score 48; DB 10; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.0043;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY      1 MXRXXXCXXXXXXXXXXXXXXXXXXPPXXXXXGXXXXXXXXXXXXXXXXXXQTXLXNEXXL 55
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MSRTIFCTYLQDAEQDFQLYPGELGKRIYNEISKDAWAQWQHKTMLINEKKL 55

RESULT 11
US-09-955-502-18
; Sequence 18, Application US/09955502
; Patent No. US20020072118A1
; GENERAL INFORMATION:
; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-09-955-502-18
```



```

; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Bordetella parapertussis
US-09-955-502-3

```

Query Match 31.7%; Score 44; DB 10; Length 87;
Best Local Similarity 26.8%; Pred. No. 0.03;
Matches 11; Conservative 0; Mismatches 30; Indels

Qy 23 PXXGXXXXXXXXXXWXXQTXLXNEXLXXXXXXXX 63

Db 23 PGELGTRINQISKEAWEWKIQTRLVNENRLNADAR 63

RESULT 31

US-09-955-502-25 ; Sequence 25, Application US/099555502
; Patent No. US20020072118A1
; GENERAL INFORMATION:

```

; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559

```

```

: CURRENT APPLICATION NUMBER: US/09/955,502
:
: CURRENT FILING DATE: 2001-09-18
:
: PRIOR APPLICATION NUMBER: 60/234,588
:
: PRIOR FILING DATE: 2000-09-22

```

```

; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 87

```

```

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-955-502-25

```

Query Match 31.7%; Score 44; DB 10; Length 87;
Best Local Similarity 30.3%; Pred. No. 0.03;
Matches 10; Conservative 0; Mismatches 23; Indels

Qy . 23 PXXXGXXXXXXXXXXXXXXXQTXLXNEXXL 55
Db 23 PGAKGEDIYNNVSRKAWEQKHQTMLINERRL 55

RESULT 32

US 09 555 502 A2
; Sequence 22, Application US/099555502
; Patent No. US20020072118A1
; GENERAL INFORMATION:

```

; APPLICANT: Downs, Diana M.
; APPLICANT: Gralnick, Jeff A.
; TITLE OF INVENTION: Method for Preventing Superoxide Damage to Cells and
; TITLE OF INVENTION: Oxygen-Labile Proteins
; FILE REFERENCE: 960296.97559

```

```

; REFERENCE: 2004200-07353
; CURRENT APPLICATION NUMBER: US/09/955,502
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: 60/234,588
; PRIOR FILING DATE: 2000-09-22

```

```

; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 22
; LENGTH: 89

```

; TYPE: PRT
 ; ORGANISM: Xylella fastidiosa
 US-09-955-502-22

Query Match 31.7%; Score 44; DB 10; Length 89;

; LOCATION: (56)..(62)
; OTHER INFORMATION: can be any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (64)..(65)
; OTHER INFORMATION: can be any amino acid
US-09-955-502-1

Query Match 30.2%; Score 42; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | | | | | | | | | | | | | |
Db 39 WXXWXXXQTXLXNEXXL 55

RESULT 34
US-10-156-761-14669
; Sequence 14669, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14669
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14669

Query Match 23.7%; Score 33; DB 15; Length 393;
Best Local Similarity 54.5%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXWXXXQTXLXNE 52
| | | | | | |
Db 220 WRVAQTTLNNE 230

RESULT 35
US-09-912-020-335
; Sequence 335, Application US/09912020
; Patent No. US20020045592A1
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; TITLE OF INVENTION: ESCHERICHIA COLI
; FILE REFERENCE: ELITRA.001DV1
; CURRENT APPLICATION NUMBER: US/09/912,020
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/492,709
; PRIOR FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: 60/117,405
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 335
; LENGTH: 99
; TYPE: PRT
; ORGANISM: E. Coli
US-09-912-020-335

Query Match 21.6%; Score 30; DB 10; Length 99;
Best Local Similarity 28.6%; Pred. No. 33;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | | | |
Db 82 WSAWRLVKTTLKQQ 95

RESULT 36
US-09-867-550-1422
; Sequence 1422, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1422
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (56)
; OTHER INFORMATION: Wherein Xaa may be any one of Arg or Gly or Trp
US-09-867-550-1422

Query Match 21.6%; Score 30; DB 10; Length 138;
Best Local Similarity 36.4%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | | | |
Db 64 WSSWRSSSTAL 74

RESULT 37
US-09-864-761-35638
; Sequence 35638, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35638
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010141.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: AA704457.1, EVALUE 5.00e-08
; OTHER INFORMATION: SWISSPROT HIT: P10263, EVALUE 9.00e-08
US-09-864-761-35638

Query Match 21.6%; Score 30; DB 10; Length 150;
Best Local Similarity 36.4%; Pred. No. 45;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | : | |
Db 29 WSVWSAIKTAL 39

RESULT 38

US-09-864-761-38059
; Sequence 38059, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38059
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF051934.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
US-09-864-761-38059

Query Match 20.9%; Score 29; DB 10; Length 58;
Best Local Similarity 30.8%; Pred. No. 35;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | : | |
Db 19 WTTWEPAATLIFN 31

RESULT 39

US-10-025-676-15
; Sequence 15, Application US/10025676
; Publication No. US2002013847A1
; GENERAL INFORMATION:
; APPLICANT: GUTIERREZ-ARMENTA, CRISANTO
; APPLICANT: SANZ-BURGOS, ANDRES P.

; APPLICANT: XIE, QI
; APPLICANT: LOPEZ, PAULA S.
; TITLE OF INVENTION: PLANT RETINOBLASTOMA-ASSOCIATED PROTEINS
; FILE REFERENCE: 604-468
; CURRENT APPLICATION NUMBER: US/10/025,676
; PRIOR FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: 09/213,294
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: PCT/ES96/00130
; PRIOR FILING DATE: 1996-06-13
; PRIOR APPLICATION NUMBER: PCT/EP97/03070
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Gallus sp.
US-10-025-676-15

Query Match 20.9%; Score 29; DB 15; Length 131;
Best Local Similarity 42.9%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNEXXL 55
| | | | |
Db 43 WTLFQHTLQNESEL 56

RESULT 40
US-09-731-872-333
; Sequence 333, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 333
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -23...-1
US-09-731-872-333

Query Match 20.9%; Score 29; DB 11; Length 150;
Best Local Similarity 29.4%; Pred. No. 73;
Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 72 WKKWKKHQKCLKKQASL 88

RESULT 41
US-09-731-872-360
; Sequence 360, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin

; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 360
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -23...-1
US-09-731-872-360

Query Match 20.9%; Score 29; DB 11; Length 177;
Best Local Similarity 29.4%; Pred. No. 83;
Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 99 WKKWKKHQKCLKKQASL 115

RESULT 42
US-10-102-806-504
; Sequence 504, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL03P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 504
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (292)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (342)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
US-10-102-806-504

Query Match 20.9%; Score 29; DB 15; Length 424;
Best Local Similarity 45.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | |
Db 338 WNAWXALETYL 348

RESULT 43
US-09-981-876-144
; Sequence 144, Application US/09981876
; Patent No. US20020164669A1
; GENERAL INFORMATION:

APPLICANT: Rosen. et al.
TITLE OF INVENTION: 70 Human Secreted Proteins
FILE REFERENCE: PZ001P1
CURRENT APPLICATION NUMBER: US/09/981,876
CURRENT FILING DATE: 2001-10-19
PRIOR APPLICATION NUMBER: 09/148,545
PRIOR FILING DATE: 1998-09-04
PRIOR APPLICATION NUMBER: 60/040,162
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/040,333
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/038,621
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/040,161
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/040,626
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/040,334
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/040,336
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/040,163
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: 60/047,615
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,600
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,597
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,502
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,633
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,583
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,617
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,618
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,503
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,592
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,581
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,584
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,500
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,587
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,492
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,598
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,613
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,582
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,596
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,612
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,632
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,601
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/043,580
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,568
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,314
PRIOR FILING DATE: 1997-04-11

PRIOR APPLICATION NUMBER: 60/043,569
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,311
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,671
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,674
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,669
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,312
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,313
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,672
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/043,315
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: 60/048,974
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/056,886
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,877
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,889
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,893
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,630
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,878
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,662
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,872
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,882
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,637
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,903
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,888
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,879
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,880
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,894
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,911
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,636
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,874
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,910
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,864
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,631
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,845
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/056,892
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 60/047,595
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/057,761
PRIOR FILING DATE: 05-Sep-1997
PRIOR APPLICATION NUMBER: 60/047,599
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: 60/047,588

```

; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,585
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,586
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,590
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,594
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,589
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,593
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,614
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,578
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/047,501
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,670
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/056,632
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,876
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,881
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,909
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,875
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,862
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,887
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,908
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/057,650
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/056,884
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 144
; LENGTH: 483

Query Match      20.9%; Score 29; DB 11; Length 483;
Best Local Similarity 35.7%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      39 WXXWXXXQTXLXNE 52
      | | | | |
Db      224 WDRWVRNQANLDKE 237

RESULT 44
US-09-981-876-209
; Sequence 209, Application US/09981876
; Patent No. US20020164669A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/981,876
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/148,545
; PRIOR FILING DATE: 1998-09-04

```

; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,669
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,312
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,313
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,672
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,315
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/056,886
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,877
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,889
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,893
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,630
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,878
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,662
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,872
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,882
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,637
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,903
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,888
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,879
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,880
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,894
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,911
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,636
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,874
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,910
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,864
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,631
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,892
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/047,595
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/057,761
; PRIOR FILING DATE: 05-Sep-1997
; PRIOR APPLICATION NUMBER: 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,588
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,585
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,586
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,590
; PRIOR FILING DATE: 1997-05-23

; PRIOR APPLICATION NUMBER: 60/047,594
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,589
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,593
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,614
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,578
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/047,501
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,670
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/056,632
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,876
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,881
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,909
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,875
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,862
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,887
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,908
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/057,650
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/056,884
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 209
; LENGTH: 483

Query Match 20.9%; Score 29; DB 11; Length 483;
Best Local Similarity 35.7%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 39 WXXWXXXQTXLXNE 52
| | | | |
Db. 224 WDRWVRNQANLDKE 237

RESULT 45
US-09-148-545-144
; Sequence 144, Application US/09148545
; Publication No. US20030027132A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161

EARLIER	APPLICATION NUMBER: 60/043,673
EARLIER	FILING DATE: 1997-04-11
EARLIER	APPLICATION NUMBER: 60/043,315
EARLIER	FILING DATE: 1997-04-11
EARLIER	APPLICATION NUMBER: 60/048,974
EARLIER	FILING DATE: 1997-06-06
EARLIER	APPLICATION NUMBER: 60/056,886
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,877
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,889
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,893
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,630
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,878
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,662
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,872
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,882
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,637
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,903
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,888
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,879
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,880
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,894
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,911
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,636
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,874
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,910
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,864
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,631
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,845
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/056,892
EARLIER	FILING DATE: 1997-08-22
EARLIER	APPLICATION NUMBER: 60/047,595
EARLIER	FILING DATE: 1997-05-23
EARLIER	APPLICATION NUMBER: 60/057,761
EARLIER	FILING DATE: 05-Sep-1997
EARLIER	APPLICATION NUMBER: 60/047,599
EARLIER	FILING DATE: 1997-05-23
EARLIER	APPLICATION NUMBER: 60/047,588
EARLIER	FILING DATE: 1997-05-23
EARLIER	APPLICATION NUMBER: 60/047,585
EARLIER	FILING DATE: 1997-05-23
EARLIER	APPLICATION NUMBER: 60/047,586
EARLIER	FILING DATE: 1997-05-23
EARLIER	APPLICATION NUMBER: 60/047,590
EARLIER	FILING DATE: 1997-05-23
EARLIER	APPLICATION NUMBER: 60/047,594
EARLIER	FILING DATE: 1997-05-23
EARLIER	APPLICATION NUMBER: 60/047,589
EARLIER	FILING DATE: 1997-05-23
EARLIER	APPLICATION NUMBER: 60/047,593
EARLIER	FILING DATE: 1997-05-23
EARLIER	APPLICATION NUMBER: 60/047,614

;
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 144
; LENGTH: 483

Query Match 20.9%; Score 29; DB 12; Length 483;
Best Local Similarity 35.7%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 224 WDRWVRNOANLDKE 237

RESULT 46

US-09-148-545-209
; Sequence 209, Application US/09148545
; Publication No. US20030027132A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07

;
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 05-Sep-1997
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 209
; LENGTH: 483

Query Match 20.9%; Score 29; DB 12; Length 483;
Best Local Similarity 35.7%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXQTXLXNE 52
| | | | |
Db 224 WDRWVRNQANLDKE 237

RESULT 47
US-09-892-949-93
; Sequence 93, Application US/09892949
; Publication No. US20030096339A1
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Presnell, Scott R.
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Kuijper, Joseph L.
; APPLICANT: Maurer, Mark F.
; TITLE OF INVENTION: CYTOKINE RECEPTOR ZCYTOR17
; FILE REFERENCE: 00-42
; CURRENT APPLICATION NUMBER: US/09/892,949
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 60/214,282
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 60/214,955
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/267,963
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-892-949-93

Query Match 20.9%; Score 29; DB 12; Length 547;
Best Local Similarity 28.6%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | : |
Db 224 WSKWSKEETRVTME 237

RESULT 48

US-09-934-455-426
; Sequence 426, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddle, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omaira
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 426
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-934-455-426

Query Match 20.9%; Score 29; DB 12; Length 639;
Best Local Similarity 35.7%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | : |
Db 342 WKGWERSQEDLKQE 355

RESULT 49

US-09-949-57
; Sequence 57, Application US/09892949
; Publication No. US20030096339A1
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Presnell, Scott R.
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Kuijper, Joseph L.
; APPLICANT: Maurer, Mark F.
; TITLE OF INVENTION: CYTOKINE RECEPTOR ZCYTOR17
; FILE REFERENCE: 00-42
; CURRENT APPLICATION NUMBER: US/09/892,949
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 60/214,282
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 60/214,955
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/267,963
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 57

..

; LENGTH: 662
; TYPE: PRT
; ORGANISM: mus musculus
US-09-892-949-57

Query Match 20.9%; Score 29; DB 12; Length 662;
Best Local Similarity 28.6%; Pred. No. 2.3e+02;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | : |
Db 224 WSKWSKEETRVTME 237

RESULT 50

US-09-972-708-12
; Sequence 12, Application US/09972708
; Publication No. US20030059871A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Cosman, David J.
; APPLICANT: Mosley, Bruce A.
; APPLICANT: Bird, Timothy A.
; APPLICANT: DuBose, Robert F.
; APPLICANT: Wiley, Steven R.
; TITLE OF INVENTION: HEMATOPOIETIN RECEPTORS HPR1 AND HPR2
; FILE REFERENCE: 3160-B
; CURRENT APPLICATION NUMBER: US/09/972,708
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-972-708-12

Query Match 20.9%; Score 29; DB 12; Length 726;
Best Local Similarity 28.6%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | : |
Db 207 WSKWSKEETRVTME 220

Search completed: July 18, 2003, 15:56:11
Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:55:14 ; Search time 40 Seconds
(without alignments)
156.218 Million cell updates/sec

Title: US-09-955-502-1
Perfect score: 139
Sequence: 1 MXXXXCXXXXXXXXXXXXX.....QTXLNEXLXXXXXXXXXX 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 2832224

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	49	35.3	90	2	AI0116	conserved hypothet	
2	48	34.5	91	2	A65082	hypothetical prote	
3	48	34.5	91	2	A85954	hypothetical prote	
4	48	34.5	91	2	AH0879	conserved hypothet	
5	48	34.5	91	2	F91108	hypothetical prote	
6	47	33.8	90	2	C82320	conserved hypothet	
7	47	33.8	90	2	C64013	hypothetical prote	
8	46	33.1	88	2	H81014	conserved hypothet	
9	46	33.1	93	2	E84994	hypothetical prote	
10	44	31.7	90	2	H83003	conserved hypothet	
11	44	31.7	105	2	C82624	conserved hypothet	
12	34	24.5	589	2	A34341	poly(3-hydroxybuty	
13	33	23.7	564	2	F90965	hypothetical prote	
14	33	23.7	564	2	F85813	hypothetical prote	
15	33	23.7	569	2	H64959	probable membrane	
16	33	23.7	1235	2	AC1358	ATP-dependent deox	
17	32	23.0	380	2	T28251	ORF MSV090 probabl	
18	32	23.0	420	2	S77102	hypothetical prote	
19	32	23.0	1194	2	C59436	KIAA1391 protein [
20	32	23.0	3655	2	T38084	TRAP-like protein	
21	31	22.3	211	2	T10392	late expression fa	
22	31	22.3	567	2	AC0754	probable membrane	
23	31	22.3	593	2	C64097	probable soluble l	
24	30	21.6	99	2	F91126	hypothetical prote	
25	30	21.6	99	2	E85971	hypothetical prote	
26	30	21.6	99	2	AE0895	conserved hypothet	
27	30	21.6	99	2	A65099	hypothetical prote	
28	30	21.6	331	1	C69026	acetylpolysamine am	
29	30	21.6	331	2	T04525	hypothetical prote	

ALIGNMENTS

RESULT 1
AI0116

conserved hypothetical protein YPO0953 [imported] - Yersinia pestis (strain CO92)
C.Species: Yersinia pestis
C.Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C.Accession: AI0116
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice,
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Douga
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Ba
Nature 413, 523-527, 2001
A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB00001; PMID:21470413; PMID:11586360

Query Match	35.3%;	Score 49;	DB 2;	Length 90;
Best Local Similarity	23.6%;	Pred. No. 0.0022;		
Matches 13;	Conservative	0;	Mismatches 42;	Indels 0;
Gaps	0;			

Qy	1	MXXXXCXXXXXXXXXXXXXXXXXXXXPXXGXXXXXXXWXXWXQTXLNEXXL	55
Db	1	MSRTIFCTFLKKDAERDQLYPGEIGRKRIYNEISKEAWSQWITKOTMLINEKKL	55

RESULT 2
A65082

hypothetical protein b2962 - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C:\Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:\Accession: A65082

R. Blattner, F.R.: Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462. 1997

A;Title: The complete genome sequence of *Escherichia coli* K-12.

A;Reference number: A64720; MUID:97426617; PMID:92785033

A; Accession: A65082

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A; Molecule type: DNA

A;Residues: 1-91 <BLAT>

A;Cross-references: GB:AE000378; GB:U00096; NID:gl789319; PIDN:AAC75999.1; PID:gl789319
A;Experimental source: strain K-12, substrain MGL1655

```

Query Match      34.5%; Score 48; DB 2; Length 91;
Best Local Similarity 23.6%; Pred. No. 0.0036;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```

QY 1 MXRXXXXXXXXXXXXXXXGXXXXXXXWXXXQTXLXNEXXL 55

Db 54 WAAWLVHQTMLINENRL 70

RESULT 12

A34341

poly(3-hydroxybutyrate) synthase (EC 2.3.1.-) - Alcaligenes eutrophus

C;Species: Alcaligenes eutrophus

C;Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 08-Sep-2000

C;Accession: A34341; A39190

R;Peoples, O.P.; Sinskey, A.J.

J. Biol. Chem. 264, 15298-15303, 1989

A;Title: Poly-beta-hydroxybutyrate (PHB) biosynthesis in Alcaligenes eutrophus H16. Ident

A;Reference number: A34341; MUID:89359357; PMID:2670936

A;Accession: A34341

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-589 <PEO>

A;Cross-references: GB:J05003; NID:gl41958; PIDN:AAA21975.1; PID:gl41959

A;Experimental source: strain H16

R;Schubert, P.; Krueger, N.; Steinbuechel, A.

J. Bacteriol. 173, 168-175, 1991

A;Title: Molecular analysis of the Alcaligenes eutrophus poly(3-hydroxybutyrate) biosynt

omoter.

A;Reference number: A39190; MUID:91100279; PMID:1987116

A;Accession: A39190

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-219 <SCH>

A;Cross-references: GB:M64341; NID:gl41964; PIDN:AAA21979.1; PID:gl41965

A;Note: the authors translated the codon TAC for residue 120 as Thr

C;Superfamily: poly(3-hydroxyalkanoic acid) synthase phbc

C;Keywords: acyltransferase

Query Match 24.5%; Score 34; DB 2; Length 589;

Best Local Similarity 42.9%; Pred. No. 17;

Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52

Db 436 WYCWYLRHTYLQNE 449

RESULT 13

F90965

hypothetical protein ECS2694 [imported] - Escherichia coli (strain O157:H7, substrain R

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C;Accession: F90965

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: F90965

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-564 <HAY>

A;Cross-references: GB:BA000007; PIDN:BAB36117.1; PID:gl3362162; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECS2694

Query Match 23.7%; Score 33; DB 2; Length 564;

Best Local Similarity 35.3%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55

Db 396 WQAWHDTLTRLYNRGAL 412

RESULT 14

F85813

hypothetical protein Z3047 [imported] - Escherichia coli (strain O157:H7, substrain

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C;Accession: F85813

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Ma-

illler, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apod-

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: F85813

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-564 <STO>

A;Cross-references: GB:AE005174; NID:gl2516062; PIDN:AAG56970.1; GSPDB:GN00145; UWGP

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: Z3047

Query Match 23.7%; Score 33; DB 2; Length 564;

Best Local Similarity 35.3%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0.

QY 39 WXXWXXXQTXLXNEXXL 55

Db 396 WQAWHDTLTRLYNRGAL 412

RESULT 15

H64959

probable membrane protein b1956 - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C;Accession: H64959

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.==

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: H64959

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-569 <BLAT>

A;Cross-references: GB:AE000287; GB:U00096; NID:gl788257; PIDN:AAC75022.1; PID:gl7882=

A;Experimental source: strain K-12, substrain MGI655

C;Genetics:

A;Start codon: TTG

C;Keywords: transmembrane protein

F;30-46/Domain: transmembrane #status predicted <TM1>

F;363-379/Domain: transmembrane #status predicted <TM2>

Query Match 23.7%; Score 33; DB 2; Length 569;

Best Local Similarity 35.3%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55

Db 401 WQAWHDTLTRLYNRGAL 417

RESULT 16

AC1358

ATP-dependent deoxyribonuclease (chain A) homolog lmo2267 [imported] - Listeria monoc=

C;Species: Listeria monocytogenes

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C;Accession: AC1358

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec=

.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1358
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1235 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00345.1; PID:g16411737; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2267

Query Match 23.7%; Score 33; DB 2; Length 1235;
Best Local Similarity 30.8%; Pred. No. 53;
Matches 4; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | : | : |
Db 276 WSSWTHLKTSIEN 288

RESULT 17

T28251
ORF MSV090 probable Molluscum contagiosum virus MC121L (vaccinia A16L) homolog - Melanop
C:Species: Melanoplus sanguinipes entomopoxvirus
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T28251
R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A:Reference number: Z20484; MUID:99102612; PMID:9847359
A:Accession: T28251
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-380 <AFO>
A:Cross-references: EMBL:AF063866; NID:g4049647; PIDN:AAC97640.1; PID:g4049680
C:Genetics:
A:Note: MSV090

Query Match 23.0%; Score 32; DB 2; Length 380;
Best Local Similarity 42.9%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 42 WXXQTXLXNEXXL 55
| | | | : |
Db 186 WLESQTKLNNDIAL 199

RESULT 18

S77102
hypothetical protein slr1865 - Synecocystis sp. (strain PCC 6803)
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S77102
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S77102
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-420 <KAN>
A:Cross-references: EMBL:D90908; GB:AB001339; NID:gl652725; PIDN:BAAL7660.1; PID:d101839
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Start codon: GTG

Query Match 23.0%; Score 32; DB 2; Length 420;
Best Local Similarity 35.7%; Pred. No. 32;
Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | : |

Db 150 WRFWEVAQTNLIKQ 163

RESULT 19

C59436
KIAA1391 protein [imported] - huam
C:Species: Homo sapiens (man)
C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 19-Jul-2002
C:Accession: C59436; D59436
R:Nagase, T.; Kikuno, R.; Ishikawa, K.I.; Hirose, M.; Ohara, O.
DNA Res. 7, 65-73, 2000
A:Title: Prediction of the coding sequences of unidentified human genes. XVI. The com
A:Reference number: C59436
A:Accession: C59436
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1194 <NAG>
A:Cross-references: GB:BAA92629; PID:g7243163; PIDN:BAA92629.1
R:Ohara, O.; Nagase, T.; Kikuno, R.
submitted to GenBank, January 2000
A:Reference number: D59436
A:Accession: D59436
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1194 <OHA>
A:Cross-references: GB:BAA92629; PID:g7243163; PIDN:BAA92629.1

Query Match 23.0%; Score 32; DB 2; Length 1194;
Best Local Similarity 38.5%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | : |
Db 340 WAFWRGSGSTHLDN 352

RESULT 20

T38084
TRAP-like protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38084
R:Gentles, S.; Odell, C.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V
submitted to the EMBL Data Library, November 1995
A:Reference number: Z21768
A:Accession: T38084
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-3655 <GEN>
A:Cross-references: EMBL:Z68136; PIDN:CAA92239.1; GSPDB:GN00066; SPDB:SPAC1F5.11c
A:Experimental source: strain 972h; cosmid clF5
C:Genetics:
A:Gene: SPDB:SPAC1F5.11c
A:Map position: 1
A:Introns: 22/1

Query Match 23.0%; Score 32; DB 2; Length 3655;
Best Local Similarity 54.5%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXQTXLXNE 52
| | | | : |
Db 1553 WLFFQTILTNE 1563

RESULT 21

T10392
late expression factor 7 protein - Orgyia pseudotsugata nuclear polyhedrosis virus
C:Species: Orgyia pseudotsugata nuclear polyhedrosis virus, OpMNPV
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10392
R:Ahrens, C.A.; Russell, R.R.; Funk, C.J.; Evans, J.; Harwood, S.; Rohrmann, G.F.
Virology 229, 381-399, 1997

A;Title: The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear polyhedrosis
A;Reference number: Z17011; MUID:97271300; PMID:9126251
A;Accession: T10392
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-211 <AHR>
A;Cross-references: EMBL:U75930; NID:g2934903; PIDN:AAC59122.1; PID:g1911369
C;Superfamily: Orgyia pseudotsugata nuclear polyhedrosis virus late expression factor 7

Query Match 22.3%; Score 31; DB 2; Length 211;
Best Local Similarity 35.7%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | | : |
Db 133 WNWGGLTRLLIHE 146

RESULT 22
AC0754
probable membrane protein STY2194 [imported] - Salmonella enterica serov
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C;Accession: AC0754
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; PMID:11677608
A;Accession: AC0754
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-567 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05734.1; PID:g16503227; GSPDB:GN00176
C;Genetics:
A;Gene: STY2194

Query Match 22.3%; Score 31; DB 2; Length 567;
Best Local Similarity 35.3%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | : | | : |
Db 393 WQAWHDPLTRLNREGAL 409

RESULT 23
C64097
probable soluble lytic transglycosylase (EC 3.2.1.-) - Haemophilus influenzae (strain Rd
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 22-Jun-1999
C;Accession: C64097
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: C64097
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-593 <TIGR>
A;Cross-references: GB:U32765; GB:L42023; NID:g1573838; PIDN:AAC22487.1; PID:g1573843; T
C;Superfamily: soluble lytic transglycosylase
C;Keywords: glycosidase; hydrolase

Query Match 22.3%; Score 31; DB 2; Length 593;
Best Local Similarity 35.3%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | : | | : |
Db 300 WRYWEAKQDILKNTKKL 316

RESULT 24
F91126
hypothetical protein ECS3982 [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: F91126
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han,
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F91126
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-99 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA37405.1; PID:g13363455; GSPDB:GN0015
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECS3982

Query Match 21.6%; Score 30; DB 2; Length 99;
Best Local Similarity 28.6%; Pred. No. 23;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | | : |
Db 82 WSAWRLVKTTLKQQ 95

RESULT 25
E85971
hypothetical protein Z4454 [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: E85971
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Ma
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apod
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85971
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-99 <STO>
A;Cross-references: GB:AE005174; NID:g12517690; PIDN:AAG58233.1; GSPDB:GN0014
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z4454

Query Match 21.6%; Score 30; DB 2; Length 99;
Best Local Similarity 28.6%; Pred. No. 23;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | | : |
Db 82 WSAWRLVKTTLKQQ 95

RESULT 26
AE0895
conserved hypothetical protein STY3411 [imported] - Salmonella enterica subsp. enter
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C;Accession: AE0895
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Far

S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AE0895
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-99 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD07755.1; PID:g16504306; GSPDB:GN00176
C:Genetics:
A:Gene: STY3411

Query Match 21.6%; Score 30; DB 2; Length 99;
Best Local Similarity 28.6%; Pred. No. 23;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | | :
Db 82 WSAWRLVKTTLRQQ 95

RESULT 27
A65099
hypothetical protein b3100 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: A65099
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: A65099
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-99 <BLAT>
A:Cross-references: GB:AE000392; GB:U00096; NID:g2367194; PIDN:AAC76135.1; PID:g1789487;
A:Experimental source: strain K-12, substrain MGL655

Query Match 21.6%; Score 30; DB 2; Length 99;
Best Local Similarity 28.6%; Pred. No. 23;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | | :
Db 82 WSAWRLVKTTLRQQ 95

RESULT 28
C69026
acetylpolylamine aminohydrolase - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: C69026
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: C69026
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-331 <MTH>
A:Cross-references: GB:AE000887; GB:AE000666; NID:g2622289; PIDN:AAB85683.1; PID:g262230
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1194
C:Superfamily: acetylpolylamine aminohydrolase; RPD3/acuC homology
F:3-298/Domain: RPD3/acuC homology <RAH1>

Query Match 21.6%; Score 30; DB 1; Length 331;

Best Local Similarity 45.5%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXL 49
| | : | | :
Db 263 WIGWEIHTGTGL 273

RESULT 29
T04525
hypothetical protein F16A16.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 04-Mar-2000
C:Accession: T04525
R:Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Hoheisel, J.; M
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15376
A:Accession: T04525
A:Molecule type: DNA
A:Residues: 1-331 <BEV>
A:Cross-references: EMBL:AL035353
A:Experimental source: cultivar Columbia; BAC clone F16A16
C:Genetics:
A:Map position: 4
A:Introns: 76/2; 269/2; 286/2
A:Note: F16A16.150
C:Superfamily: Arabidopsis thaliana hypothetical protein F16A16.150

Query Match 21.6%; Score 30; DB 2; Length 331;
Best Local Similarity 38.5%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | : | | :
Db 267 WEKWLNEQKKLAN 279

RESULT 30
A86227
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86227
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, R.; Rizzio,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Mdrzja
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86227
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-385 <STO>
A:Cross-references: GB:AE005172; NID:g3482913; PIDN:AAC33198.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 21.6%; Score 30; DB 2; Length 385;
Best Local Similarity 36.4%; Pred. No. 78;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | : | | :
Db 208 WAAWFLIQTGM 218

RESULT 31
T39230

hypothetical protein SPAC9G1.07 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T39230
R:Churcher, C.M.; Gentles, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21837
A:Accession: T39230
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-418 <CHU>
A:Cross-references: EMBL:Z98763; PIDN:CAB11491.1; GSPDB:GN00066; SPDB:SPAC9G1.07
A:Experimental source: strain 972h-; cosmid c9G1
C:Genetics:
A:Gene: SPDB:SPAC9G1.07
A:Map position: 1

Query Match 21.6%; Score 30; DB 2; Length 418;
Best Local Similarity 29.4%; Pred. No. 84;
Matches 5; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 319 WGSWVSQDTSKNSSNL 335

RESULT 32
TVCH2
transcription factor ets-2 - chicken
C:Species: Gallus gallus (chicken)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C:Accession: S00386
R:Boulukos, K.E.; Pognonec, P.; Begue, A.; Galibert, F.; Gesquiere, J.C.; Stehelin, D.;
EMBO J. 7, 697-705, 1988
A:Title: Identification in chickens of an evolutionarily conserved cellular ets-2 gene
A:Reference number: S00386; MUID:88283637; PMID:3293999
A:Accession: S00386
A:Molecule type: mRNA
A:Residues: 1-479 <BOU>
A:Cross-references: EMBL:X07202; NID:G63181; PIDN:CAA30178.1; PID:G63182
A:Note: it is uncertain whether 1-Met, 9-Met, 18-Met, or 22-Met is the initiator codon
C:Genetics:
A:Gene: ets-2
C:Superfamily: transcription factor ets; ets DNA-binding domain homology; ets RII regula
C:Keywords: DNA binding; nucleus; phosphoprotein; proto-oncogene; transcription factor;
F;94-167/Domain: ets RII regulatory region homology <ETR>
F;134-167/Region: helix-loop-helix #status predicted
F;375-453/Domain: ets DNA-binding domain homology <ETS>
F;415-421/Region: nuclear location signal

Query Match 21.6%; Score 30; DB 1; Length 479;
Best Local Similarity 30.8%; Pred. No. 95;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 117 WLAWATNEFSLAN 129

RESULT 33
C70606
probable fadE34 protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: C70606
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70606

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-711 <COL>
A:Cross-references: GB:Z92774; GB:AL123456; NID:G3261729; PIDN:CAB07147.1; PID:e3067
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: fadE34

Query Match 21.6%; Score 30; DB 2; Length 711;
Best Local Similarity 45.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNE 52
| | | | |
Db 571 WRLARTLANE 581

RESULT 34
T29644
hypothetical protein W01A11.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29644
R:Blanchard, M.; Bradshaw, H.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid W01A11.
A:Reference number: Z20658
A:Accession: T29644
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-823 <BLA>
A:Cross-references: EMBL:U64852; PIDN:AAB04966.1; GSPDB:GN00023; CESP:W01A11.3
A:Experimental source: strain Bristol N2; clone W01A11
C:Genetics:
A:Gene: CESP:W01A11.3
A:Map position: 5
A:Introns: 15/3; 87/2; 137/2; 207/1; 258/3; 286/2; 419/3; 594/2; 729/2

Query Match 21.6%; Score 30; DB 2; Length 823;
Best Local Similarity 30.8%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 641 WDAWNSRQNDIRN 653

RESULT 35
B75573
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: B75573
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-546 <WHI>
A:Cross-references: GB:AE001863; GB:AE001825; NID:G6460670; PIDN:AAF12422.1; PID:G646

A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0207
A:Map position: 2

Query Match 21.2%; Score 29.5; DB 2; Length 546;
Best Local Similarity 38.9%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 39 WXXWXXXQ-TXLXNEXXL 55
| | | | |
Db 281 WSAWRRQQVTNLVRETAL 298

RESULT 36
S76185
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C;Accession: S76185
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76185
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-124 <KAN>
A;Cross-references: EMBL:D90914; GB:AB001339; NID:gl653477; PIDN:BAA18444.1; PID:dl01917
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 20.9%; Score 29; DB 2; Length 124;
Best Local Similarity 28.6%; Pred. No. 45;
Matches 4; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | | :
Db 97 WLSWLHGETNLIHK 110

RESULT 37
JC5854
polyketide synthase (EC 2.-.-.-) chain 5 - Actinomadura hibisca
C;Species: Actinomadura hibisca
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 17-Mar-1999
C;Accession: JC5854
R;Dairi, T.; Hamano, Y.; Igarashi, Y.; Furumai, T.; Oki, T.
Biosci. Biotechnol. Biochem. 61, 1445-1453, 1997
A;Title: Cloning and nucleotide sequence of the putative polyketide synthase genes for P
A;Reference number: JC5850; MUID:97480928; PMID:9339544
A;Accession: JC5854
A;Molecule type: DNA
A;Residues: 1-153 <DAI>
A;Cross-references: DDBJ:D87924
C;Comment: This enzyme catalyzes repeated condensation cycles of acyl-CoA, affecting the
C;Genetics:
A;Gene: pms5
C;Keywords: transferase

Query Match 20.9%; Score 29; DB 2; Length 153;
Best Local Similarity 36.4%; Pred. No. 55;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | : | |
Db 127 WAGWNAARTGL 137

RESULT 38
QJECAL
cell division inhibitor sula - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 24-Sep-1981 #sequence_revision 15-Oct-1996 #text_change 01-Mar-2002
C;Accession: A29016; A04468; I57720; E64836
R;Freudl, R.; Braun, G.; Honore, N.; Cole, S.T.
Gene 52, 31-40, 1987
A;Title: Evolution of the enterobacterial sula gene: a component of the SOS system encod
A;Reference number: A29016; MUID:87248093; PMID:3297925
A;Accession: A29016

A;Molecule type: DNA
A;Residues: 1-169 <FRE>
R;Beck, E.; Bremer, E.
Nucleic Acids Res. 8, 3011-3024, 1980
A;Title: Nucleotide sequence of the gene ompA coding the outer membrane protein II of
A;Reference number: A93707; MUID:81053729; PMID:6253901
A;Accession: A04468
A;Molecule type: DNA
A;Residues: 1-144, 'IHSGKRILSRHETTFRAKNSL' <BEC>
A;Experimental source: strain K-12
A;Note: this sequence has been revised by personal communication to the authors of re
R;Cole, S.T.
Mol. Gen. Genet. 189, 400-404, 1983
A;Title: Characterisation of the promoter for the LexA regulated sula gene of Escheri
A;Reference number: I57720; MUID:83244178; PMID:6306396
A;Accession: I57720
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10 <RES>
A;Cross-references: EMBL:V00358; NID:g43023; PIDN:CAA23654.1; PID:g43024
A;Note: the sequence represented here from this article is quoted from references A29
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: E64836
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-169 <BLAT>
A;Cross-references: GB:AE000198; GB:U00096; NID:gl787189; PIDN:AAC74044.1; PID:gl7871
A;Experimental source: strain K-12, substrain MGL655
C;Genetics:
A;Gene: sula; sfiA
A;Map position: 22 min
C;Function:
A;Description: inducible cell division inhibitor; component of the SOS response; bloc
x which may be the mechanism by which sula inhibits cell division
A;Note: expression of sula is repressed by lexA protein
C;Superfamily: cell division inhibitor sula
C;Keywords: cell division control; SOS response

Query Match 20.9%; Score 29; DB 1; Length 169;
Best Local Similarity 35.7%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 63 WQLWLTPQKLSRE 76

RESULT 39
B29016
cell division inhibitor sula - Salmonella typhimurium
C;Species: Salmonella typhimurium
C;Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 01-Feb-2002
C;Accession: B29016
R;Freudl, R.; Braun, G.; Honore, N.; Cole, S.T.
Gene 52, 31-40, 1987
A;Title: Evolution of the enterobacterial sula gene: a component of the SOS system en
A;Reference number: A29016; MUID:87248093; PMID:3297925
A;Accession: B29016
A;Molecule type: DNA
A;Residues: 1-169 <FRE>
A;Cross-references: GB:M16324; NID:gl54377; PIDN:AAA27230.1; PID:gl54378
C;Genetics:
A;Gene: sula
C;Function:
A;Description: inducible cell division inhibitor
A;Note: expression of sula is repressed by lexA protein
C;Superfamily: cell division inhibitor sula
C;Keywords: cell division control; SOS response

Query Match 20.9%; Score 29; DB 2; Length 169;
Best Local Similarity 35.7%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 63 WQLWLTPQOKLSRE 76

RESULT 40
C29016
cell division inhibitor sula - Enterobacter aerogenes
C;Species: Enterobacter aerogenes
C;Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 01-Feb-2002
C;Accession: C29016
R;Freundl, R.; Braun, G.; Honore, N.; Cole, S.T.
Gene 52, 31-40, 1987
A;Title: Evolution of the enterobacterial sula gene: a component of the SOS system encoded by the enterobacterial sula gene
A;Reference number: A29016; MUID:87248093; PMID:3297925
A;Accession: C29016
A;Molecule type: DNA
A;Residues: 1-169 <FRE>
A;Cross-references: GB:M16467; NID:gl48376; PIDN:AAA24810.1; PID:gl48377
C;Genetics:
A;Gene: sula
C;Function:
A;Description: inducible cell division inhibitor
A;Note: expression of sula is repressed by lexA protein
C;Superfamily: cell division inhibitor sula
C;Keywords: cell division control; SOS response

Query Match 20.9%; Score 29; DB 2; Length 169;
Best Local Similarity 35.7%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 63 WQLWLTPQOKLSKE 76

RESULT 41
B90759
suppressor of lon [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: B90759
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; gasawara, N.; yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic map
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B90759
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-169 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB34465.1; PID:gl3360502; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1042
C;Superfamily: cell division inhibitor sula

Query Match 20.9%; Score 29; DB 2; Length 169;
Best Local Similarity 35.7%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 63 WQLWLTPQOKLSRE 76

RESULT 42
H85622
hypothetical protein sula [imported] - Escherichia coli (strain O157:H7, substrain EDL93

C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: H85622
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Ma-
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoc-
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: H85622
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-169 <STO>
A;Cross-references: GB:AE005174; NID:gl2514143; PIDN:AAG55444.1; GSPDB:GN00145; UWGP-
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: sula
C;Superfamily: cell division inhibitor sula

Query Match 20.9%; Score 29; DB 2; Length 169;
Best Local Similarity 35.7%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 63 WQLWLTPQOKLSRE 76

RESULT 43
AB0627
cell division inhibitor [imported] - Salmonella enterica subsp. enterica serovar Typh
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C;Accession: AB0627
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church-
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Far-
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica s-
A;Reference number: AB0502; PMID:11677508
A;Accession: AB0627
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-169 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD08197.1; PID:gl6502246; GSPDB:GN00176
C;Genetics:
A;Gene: STY1092
C;Superfamily: cell division inhibitor sula

Query Match 20.9%; Score 29; DB 2; Length 169;
Best Local Similarity 35.7%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 63 WQLWLTPQOKLSRE 76

RESULT 44
AH0821
probable membrane protein STY2766 [imported] - Salmonella enterica subsp. enterica s-
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C;Accession: AH0821
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church-
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Far-
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica s-
A;Reference number: AB0502; PMID:11677608

A:Accession: AH0821
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-206 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02724.1; PID:gl6503736; GSPDB:GN00176
C:Genetics:
A:Gene: STY2766
C:Superfamily: hypothetical protein HI0370

Query Match 20.9%; Score 29; DB 2; Length 206;
Best Local Similarity 44.4%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
| | | |
Db 39 WRYWSSHQT 47

RESULT 45
H75618
cob(I)alamin adenosyltransferase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: H75618
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75618
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-207 <WHI>
A:Cross-references: GB:AE001826; NID:g6460827; PIDN:AAF12570.1; PID:g6460866; TIGR:DRB00
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRB0008
A:Map position: megaplasmid
A:Genome: plasmid
A>Note: plasmid MP1
C:Superfamily: Escherichia coli cob(I)alamin adenosyltransferase coba

Query Match 20.9%; Score 29; DB 2; Length 207;
Best Local Similarity 29.4%; Pred. No. 72;
Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLNEXXL 55
| | : | | |
Db 148 WVAWAEVPLKNRDPL 164

RESULT 46
T16103
hypothetical protein F19C7.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000
C:Accession: T16103
R;Le, T.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid F19C7.
A:Reference number: Z18461
A:Accession: T16103
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-310 <LET>
A:Cross-references: EMBL:U42439; NID:g1123106; PID:g1123109; PIDN:AAA83507.1; CESP:F19C7
C:Genetics:
A:Gene: CESP:F19C7.3
A:Introns: 71/2; 178/3; 269/2
C:Superfamily: Caenorhabditis elegans hypothetical protein C24H12.11

Query Match 20.9%; Score 29; DB 2; Length 310;

Best Local Similarity 38.5%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 66 WIEWFYNTTKLDN 78

RESULT 47
JC1069
nonstructure protein - rice dwarf virus
C:Species: rice dwarf virus
C:Date: 09-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 20-Sep-1999
C:Accession: JC1069
R;Chu, R.Y.; Zhang, X.; Pan, N.S.; Chen, Z.L.
Acta Bot. Sin. 35, 115-120, 1993
A:Title: The cDNA cloning and nucleotide sequence of the gene encoding nonstructure p
A:Reference number: JC1069
A:Accession: JC1069
A:Molecule type: mRNA
A:Residues: 1-372 <CHU>
C:Genetics:
A:Map position: segment 10
C:Superfamily: wound tumor virus nonstructural protein Pns11

Query Match 20.9%; Score 29; DB 2; Length 372;
Best Local Similarity 23.5%; Pred. No. 1.2e+02;
Matches 4; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLNEXXL 55
| | : | | |
Db 280 WRGWASTYAYMFNQEQ 296

RESULT 48
AG2787
long-chain fatty acid transport protein fadL [imported] - Agrobacterium tumefaciens (C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AG2787
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woc
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens
A:Reference number: AB2577; PMID:11743193
A:Accession: AG2787
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-425 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL42717.1; PID:gl7740155; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: fadL
A:Map position: circular chromosome

Query Match 20.9%; Score 29; DB 2; Length 425;
Best Local Similarity 44.4%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
| | | | |
Db 282 WTDWSQIQT 290

RESULT 49
A97567
hypothetical protein AGR_C_3154 [imported] - Agrobacterium tumefaciens (strain C58, C
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: A97567

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A;Reference number: A97359; PMID:11743194
A;Accession: A97567
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-429 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK87490.1; PID:g15156815; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_3154
A;Map position: circular chromosome

Query Match 20.9%; Score 29; DB 2; Length 429;
Best Local Similarity 44.4%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
| | |
Db 286 WTDWSQIQT 294

RESULT 50
S76837

phosphopyruvate hydratase (EC 4.2.1.11) - Synechocystis sp. (strain PCC 6803)
N;Alternate names: protein slr0752
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S76837
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76837
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-432 <KAN>
A;Cross-references: EMBL:D90917; GB:AB001339; NID:g1653836; PIDN:BAA18749.1; PID:g165383
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: eno
C;Superfamily: enolase
C;Keywords: carbon-oxygen lyase; gluconeogenesis; glycolysis; hydro-lyase

Query Match 20.9%; Score 29; DB 2; Length 432;
Best Local Similarity 29.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXXNEXXL 55
| | | | |
Db 297 WESWKGLTTSLSGKTQL 313

Search completed: July 18, 2003, 16:00:18
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:43:18 ; Search time 17 Seconds
(without alignments)
112.499 Million cell updates/sec

Title: US-09-955-502-1
Perfect score: 139
Sequence: 1 MXXXXXXXXXXXXXXXXXXX.....QTXLXNEXLXXXXXXXXXX 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	34	24.5	589	2	US-08-756-317-5
2	29	20.9	131	4	US-09-213-293D-15
3	29	20.9	664	3	US-09-295-186-17
4	29	20.9	778	4	US-09-460-145-2
5	29	20.9	797	4	US-09-460-145-4
6	29	20.9	913	3	US-08-827-208-3
7	29	20.9	913	4	US-09-500-358-3
8	29	20.9	913	4	US-09-498-809-3
9	28	20.1	23	1	US-07-646-531D-5
10	28	20.1	23	2	US-08-488-273-5
11	28	20.1	23	4	US-09-197-770B-11
12	28	20.1	23	6	5426100-5
13	28	20.1	71	2	US-08-152-721B-28
14	28	20.1	130	4	US-09-213-293D-17
15	28	20.1	131	4	US-09-213-293D-16
16	28	20.1	178	2	US-08-874-832-18
17	28	20.1	178	3	US-09-097-233-18
18	28	20.1	209	2	US-08-874-832-9
19	28	20.1	209	3	US-09-097-233-9
20	28	20.1	218	3	US-08-985-526-1
21	28	20.1	239	5	PCT-US93-01652-1
22	28	20.1	441	3	US-08-985-526-3
23	28	20.1	468	4	US-09-092-636-9
24	28	20.1	470	4	US-09-092-636-4
25	28	20.1	649	3	US-09-295-186-18
26	28	20.1	816	1	US-08-038-760-3
27	28	20.1	816	2	US-08-470-091-3

28	20.1	28	1	US-08-204-329-1	Sequence 1, Appli
29	20.1	28	2	US-08-959-638-8	Sequence 8, Appli
30	20.1	28	2	US-08-482-627-5	Sequence 5, Appli
31	20.1	28	3	US-08-801-092-4	Sequence 4, Appli
32	20.1	28	4	US-08-328-673A-8	Sequence 8, Appli
33	20.1	28	4	US-09-315-113-4	Sequence 4, Appli
34	20.1	28	5	PCT-US94-10357-2	Sequence 2, Appli
35	20.1	28	5	PCT-US94-10357-3	Sequence 3, Appli
36	20.1	28	4	US-09-302-812-6	Sequence 6, Appli
37	20.1	28	4	US-09-511-477-6	Sequence 6, Appli
38	20.1	28	4	US-09-511-507-6	Sequence 6, Appli
39	20.1	28	4	US-09-302-812-4	Sequence 4, Appli
40	20.1	28	4	US-09-511-477-4	Sequence 4, Appli
41	20.1	28	4	US-09-511-507-4	Sequence 4, Appli
42	20.1	28	1	US-08-313-288B-20	Sequence 20, Appli
43	19.4	27	4	US-09-342-653-7	Sequence 7, Appli
44	19.4	27	2	US-08-850-880-2	Sequence 2, Appli
45	19.4	27	2	US-08-944-916-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-756-317-5
; Sequence 5, Application US/08756317
; Patent No. 5849894
; GENERAL INFORMATION:
; APPLICANT: Clemente, Thomas E.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Mitsky, Timothy A.
; APPLICANT: Stark, David M.
; TITLE OF INVENTION: Improved Rhodospirillum Rubrum
; TITLE OF INVENTION: Poly-B-Hydroxyalkanoate Synthase
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,317
; FILING DATE: 25-NOV-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/007,693
; FILING DATE: 29-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Patterson, Melinda L.
; REGISTRATION NUMBER: 33,062
; REFERENCE/DOCKET NUMBER: MOBT:008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 787-1400
; TELEFAX: (713) 787-1440
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-756-317-5

Query Match 24.5%; Score 34; DB 2; Length 589;
Best Local Similarity 42.9%; Pred. No. 5.6;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
 Db 436 WYCWLRLRHTYLQNE 449

RESULT 2

US-09-213-293D-15
 ; Sequence 15, Application US/09213293D
 ; Patent No. 6384299
 ; GENERAL INFORMATION:
 ; APPLICANT: GUTIERREZ-ARMENTA, CRISANTO
 ; APPLICANT: SANZ-BURGOS, ANDRES P.
 ; APPLICANT: XIE, QI
 ; APPLICANT: LOPEZ, PAULA S.
 ; TITLE OF INVENTION: PLANT RETINOBLASTOMA-ASSOCIATED PROTEINS
 ; FILE REFERENCE: 604-469
 ; CURRENT APPLICATION NUMBER: US/09/213,293D
 ; CURRENT FILING DATE: 1998-12-14
 ; PRIOR APPLICATION NUMBER: PCT/ES96/00130
 ; PRIOR FILING DATE: 1996-06-13
 ; PRIOR APPLICATION NUMBER: PCT/EP97/03070
 ; PRIOR FILING DATE: 1997-06-12
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 15
 ; LENGTH: 131
 ; TYPE: PRT
 ; ORGANISM: Gallus sp.
 US-09-213-293D-15

Query Match 20.9%; Score 29; DB 4; Length 131;
 Best Local Similarity 42.9%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 42 WXXWXXXQTXLXNEXXL 55
 Db 43 WTLFQHTLQNESEL 56

RESULT 3

US-09-295-186-17
 ; Sequence 17, Application US/09295186B
 ; Patent No. 6127137
 ; GENERAL INFORMATION:
 ; APPLICANT: Hasida, Miyoko
 ; APPLICANT: Tsutsumi, No. 6127137iko
 ; APPLICANT: Halkier, Torben
 ; APPLICANT: Stringer, Mary Ann
 ; TITLE OF INVENTION: An Acidic Phospholipase, Production, and
 ; TITLE OF INVENTION: Methods of Using Thereof (As Amended)
 ; FILE REFERENCE: 4953.204-US
 ; CURRENT APPLICATION NUMBER: US/09/295,186B
 ; CURRENT FILING DATE: 1999-04-20
 ; PRIOR APPLICATION NUMBER: 1215/96
 ; PRIOR FILING DATE: 1996-10-31
 ; PRIOR APPLICATION NUMBER: PCT/DK97/00490
 ; PRIOR FILING DATE: 1997-10-30
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 17
 ; LENGTH: 664
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-09-295-186-17

Query Match 20.9%; Score 29; DB 3; Length 664;
 Best Local Similarity 30.8%; Pred. No. 85;
 Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
 Db 158 WNNWTSVQAIVDN 170

RESULT 4
 US-09-460-145-2
 ; Sequence 2, Application US/09460145
 ; Patent No. 6287838
 ; GENERAL INFORMATION:
 ; APPLICANT: Kriz, Ron
 ; APPLICANT: Song, Chuanzheng
 ; TITLE OF INVENTION: CYTOSOLIC PHOSPHOLIPASE A2-BETA ENZYMES
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/460,145
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/788,975
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brown, Scott A.
 ; REGISTRATION NUMBER: 32,724
 ; REFERENCE/DOCKET NUMBER: GI5289
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 498-8224
 ; TELEFAX: (617) 876-5851
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 778 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-460-145-2

Query Match 20.9%; Score 29; DB 4; Length 778;
 Best Local Similarity 35.7%; Pred. No. 99;
 Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
 Db 520 WDRWVRNQANLDKE 533

RESULT 5

US-09-460-145-4
 ; Sequence 4, Application US/09460145
 ; Patent No. 6287838
 ; GENERAL INFORMATION:
 ; APPLICANT: Kriz, Ron
 ; APPLICANT: Song, Chuanzheng
 ; TITLE OF INVENTION: CYTOSOLIC PHOSPHOLIPASE A2-BETA ENZYMES
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/460,145
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/788,975
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: G15289
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 797 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-460-145-4

```

Query Match 20.9%; Score 29; DB 4; Length 797;
 Best Local Similarity 35.7%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 539 WDRWVRNQANLDKE 552

```

RESULT 6

```

US-08-827-208-3
; Sequence 3, Application US/08827208
; Patent No. 6025178
; GENERAL INFORMATION:
; APPLICANT: Chiou, Xue-Chiou C.
; APPLICANT: Kramer, Ruth M.
; APPLICANT: Pickard, Richard T.
; APPLICANT: Sharp, John D.
; APPLICANT: Strifler, Beth A.
; TITLE OF INVENTION: HUMAN PHOSPHOLIPASE A2 AND RELATED
; TITLE OF INVENTION: NUCLEIC ACID COMPOUNDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: United States of America
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,208
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/014,608
; FILING DATE: 29-MAR-1996
; APPLICATION NUMBER: US 60/041,264
; FILING DATE: 19-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaylo, Paul J.
; REGISTRATION NUMBER: 36,808
; REFERENCE/DOCKET NUMBER: X-10610

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 276-0756
; TELEFAX: (317) 276-3861
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 913 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-827-208-3

```

Query Match 20.9%; Score 29; DB 3; Length 913;
 Best Local Similarity 35.7%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 655 WDRWVRNQANLDKE 668

```

RESULT 7

```

US-09-500-358-3
; Sequence 3, Application US/09500358
; Patent No. 6197569
; GENERAL INFORMATION:
; APPLICANT: Chiou, Xue-Chiou C.
; APPLICANT: Kramer, Ruth M.
; APPLICANT: Pickard, Richard T.
; APPLICANT: Sharp, John D.
; APPLICANT: Strifler, Beth A.
; TITLE OF INVENTION: HUMAN PHOSPHOLIPASE A2 AND RELATED
; TITLE OF INVENTION: NUCLEIC ACID COMPOUNDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: United States of America
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/500,358
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,208
; FILING DATE: 28-MAR-1997
; APPLICATION NUMBER: US 60/014,608
; FILING DATE: 29-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,264
; FILING DATE: 19-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaylo, Paul J.
; REGISTRATION NUMBER: 36,808
; REFERENCE/DOCKET NUMBER: X-10610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 276-0756
; TELEFAX: (317) 276-3861
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 913 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-500-358-3

```

Query Match 20.9%; Score 29; DB 4; Length 913;

Best Local Similarity 35.7%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 655 WDRWVRNQANLDKE 668

RESULT 8

US-09-498-809-3
; Sequence 3, Application US/09498809
; Patent No. 6242206
; GENERAL INFORMATION:
; APPLICANT: Chiou, Xue-Chiou C.
; APPLICANT: Kramer, Ruth M.
; APPLICANT: Pickard, Richard T.
; APPLICANT: Sharp, John D.
; APPLICANT: Striffler, Beth A.
; TITLE OF INVENTION: HUMAN PHOSPHOLIPASE A2 AND RELATED
; TITLE OF INVENTION: NUCLEIC ACID COMPOUNDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center,
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: United States of America
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,809
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/827,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,264
; FILING DATE: 19-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaylo, Paul J.
; REGISTRATION NUMBER: 36,808
; REFERENCE/DOCKET NUMBER: X-10610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 276-0756
; TELEFAX: (317) 276-3861
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 913 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-498-809-3

Query Match 20.9%; Score 29; DB 4; Length 913;
Best Local Similarity 35.7%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 655 WDRWVRNQANLDKE 668

RESULT 9

US-07-646-531D-5
; Sequence 5, Application US/07646531D
; Patent No. 5200397
; GENERAL INFORMATION:
; APPLICANT: Deutch, Alan Howard

; APPLICANT: Tuszynski, George Paul
; TITLE OF INVENTION: Peptide Fragments and Analogs of
; TITLE OF INVENTION: Thrombospondin
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. R. Grace & Co.-Conn.
; STREET: 7379 Route 32
; CITY: Columbia
; STATE: Maryland
; COUNTRY: USA
; ZIP: 21044
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/646,531D
; FILING DATE: 19910131
; ATTORNEY/AGENT INFORMATION:
; NAME: Appleby, Vanessa L.
; REGISTRATION NUMBER: 33223
; REFERENCE/DOCKET NUMBER: 01-7896
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 531-4515
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-646-531D-5

Query Match 20.1%; Score 28; DB 1; Length 23;
Best Local Similarity 30.8%; Pred. No. 6.1;
Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 4 WSEWTSCTSCGN 16

RESULT 10

US-08-488-273-5
; Sequence 5, Application US/08488273
; Patent No. 5840692
; GENERAL INFORMATION:
; APPLICANT: Deutch, Alan H.
; APPLICANT: Tuszynski, George P.
; TITLE OF INVENTION: PEPTIDE FRAGMENTS AND ANALOGS OF
; TITLE OF INVENTION: THROMBOSPONDIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103-2398
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,273
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: .514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/359,263
; FILING DATE: 19-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/110,146

;; FILING DATE: 20-AUG-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/896,527
;; FILING DATE: 09-JUN-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/483,527
;; FILING DATE: 22-FEB-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Leary Ph.D., Kathryn
;; REGISTRATION NUMBER: 36,317
;; REFERENCE/DOCKET NUMBER: 9598-204
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (215) 567-2020
;; TELEFAX: (215) 567-2991
;; TELEX: 831-494
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 23 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-488-273-5

Query Match 20.1%; Score 28; DB 2; Length 23;
Best Local Similarity 30.8%; Pred. No. 6.1;
Matches 4; Conservative 0; Mismatches 9; Indels 9; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 4 WSEWTSCSTSCGN 16

RESULT 11
US-09-197-770B-11
; Sequence 11, Application US/09197770B
; Patent No. 6339062
; GENERAL INFORMATION:
; APPLICANT: Tuszyński, George
; APPLICANT: Williams, Taffy
; APPLICANT: Actor, Paul
; TITLE OF INVENTION: RETROINVERSO POLYPEPTIDES THAT MIMIC OR INHIBIT
; TITLE OF INVENTION: THROMBOSPONDIN ACTIVITY
; FILE REFERENCE: 07206-0021
; CURRENT APPLICATION NUMBER: US/09/197,770B
; CURRENT FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: fragment/ analog of thrombospondin
US-09-197-770B-11

Query Match 20.1%; Score 28; DB 4; Length 23;
Best Local Similarity 30.8%; Pred. No. 6.1;
Matches 4; Conservative 0; Mismatches 9; Indels 9; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 4 WSEWTSCSTSCGN 16

RESULT 12
5426100-5
; Patent No. 5426100
; APPLICANT: DEUTCH, ALAN H.; TUSZYNSKI, GEORGE
; TITLE OF INVENTION: PEPTIDE FRAGMENTS AND ANALOGS OF
; THROMBOSPONDIN
; NUMBER OF SEQUENCES: 12

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/110,146
;; FILING DATE: 20-AUG-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 896,527
;; FILING DATE: 09-JUN-1992
;; APPLICATION NUMBER: 483,527
;; FILING DATE: 22-FEB-1990
;; SEQ ID NO: 5:
;; LENGTH: 23
5426100-5

Query Match 20.1%; Score 28; DB 6; Length 23;
Best Local Similarity 30.8%; Pred. No. 6.1;
Matches 4; Conservative 0; Mismatches 9; Indels 9; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 4 WSEWTSCSTSCGN 16

RESULT 13
US-08-152-721B-28
; Sequence 28, Application US/08152721B
; Patent No. 5962315
; GENERAL INFORMATION:
; APPLICANT: Livingston, David M.
; APPLICANT: Ewen, Mark E.
; TITLE OF INVENTION: DNA Encoding p107 Tumor Suppressor and
; TITLE OF INVENTION: Related Polypeptides
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHOATE, HALL & STEWART
; STREET: 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2891
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,721B
; FILING DATE: 15-NOV-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasternack Esq., Sam
; REGISTRATION NUMBER: 29,576
; REFERENCE/DOCKET NUMBER: 181411-011DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-5000
; TELEFAX: (617) 248-4000
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; IMMEDIATE SOURCE:
; CLONE: RB region 6
US-08-152-721B-28

Query Match 20.1%; Score 28; DB 2; Length 71;
Best Local Similarity 42.9%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 8; Indels 8; Gaps 0;

QY 42 WXXWXXXQTXLXNEXXL 55

STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/097,233
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 2261-1-001 N
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 178 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-097-233-18

Query Match 20.1%; Score 28; DB 3; Length 178;
Best Local Similarity 44.4%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
| | | |
Db 8 WTFWYDIQT 16

RESULT 18

US-08-874-832-9
Sequence 9, Application US/08874832
Patent No. 5872011
GENERAL INFORMATION:
APPLICANT: Burly, Stephen K.
APPLICANT: Sonenberg, Nahum
APPLICANT: Marcotrigiano, Joseph
APPLICANT: Gingras, Anne-Claude
TITLE OF INVENTION: A CRYSTAL OF A PROTEIN-LIGAND COMPLEX
TITLE OF INVENTION: CONTAINING AN N-TERMINAL TRUNCATED EIF4E, AND METHODS OF
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/874,832
FILING DATE: 13-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 2261-1-001
TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-874-832-9

Query Match 20.1%; Score 28; DB 2; Length 209;
Best Local Similarity 44.4%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
| | | |
Db 39 WTFWYDIQT 47

RESULT 19

US-09-097-233-9
Sequence 9, Application US/09097233
Patent No. 6020162
GENERAL INFORMATION:
APPLICANT: Burley, Stephen K.
APPLICANT: Sonenberg, Nahum
APPLICANT: Marcotrigiano, Joseph
APPLICANT: Gingras, Anne-Claude
TITLE OF INVENTION: A CRYSTAL OF A PROTEIN-LIGAND COMPLEX
TITLE OF INVENTION: CONTAINING AN N-TERMINAL TRUNCATED EIF4E, AND METHODS OF
TITLE OF INVENTION: USE THEREOF
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/097,233
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 2261-1-001 N
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-097-233-9

Query Match 20.1%; Score 28; DB 3; Length 209;
Best Local Similarity 44.4%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47

Db 39 WTFWYDIQT 47

RESULT 20

US-08-985-526-1

; Sequence 1, Application US/08985526

; Patent No. 6080728

; GENERAL INFORMATION:

; APPLICANT: Mixson, James A

; TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA

; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE

; TITLE OF INVENTION: THERAPY

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Connolly, Bove, Lodge, & Hutz

; STREET: 1220 Market Street, P.O. Box 2207

; CITY: Wilmington

; STATE: Delaware

; COUNTRY: U.S.A.

; ZIP: 19899

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/985,526

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/608,845

; FILING DATE: 16-JUL-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: McMorriow Jr., Robert G

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (302) 658-9141

; TELEFAX: (302) 658-5613

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 218 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-08-985-526-1

Query Match 20.1%; Score 28; DB 3; Length 218;

Best Local Similarity 30.8%; Pred. No. 51;

Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51

Db 85 WSEWTSCTSCGN 97

RESULT 21

PCT-US93-01652-1

; Sequence 1, Application PC/TUS9301652

; GENERAL INFORMATION:

; APPLICANT: Bouck, Noel P.

; APPLICANT: Polverini, Peter J.

; APPLICANT: Good, Deborah J.

; APPLICANT: Frazier, William A.

; TITLE OF INVENTION: Method and Composition for

; TITLE OF INVENTION: Inhibiting Angiogenesis

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut

; STREET: 100 South Wacker Drive, Suite 960

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606-4002

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/01652

FILING DATE: 19930222

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/841,656

FILING DATE: 24-FEB-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/464,369

FILING DATE: 12-JAN-1990

ATTORNEY/AGENT INFORMATION:

NAME: Fentress, Susan B.

REGISTRATION NUMBER: 31,327

REFERENCE/DOCKET NUMBER: 92005-PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312)-456-8000

TELEFAX: (312)-456-7776

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 239 amino acids

TYPE: AMINO ACID

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: peptide

PCT-US93-01652-1

Query Match 20.1%; Score 28; DB 5; Length 239;

Best Local Similarity 30.8%; Pred. No. 55;

Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51

Db 74 WSEWTSCTSCGN 86

RESULT 22

US-08-985-526-3

; Sequence 3, Application US/08985526

; Patent No. 6080728

; GENERAL INFORMATION:

; APPLICANT: Mixson, James A

; TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA

; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE

; TITLE OF INVENTION: THERAPY

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Connolly, Bove, Lodge, & Hutz

; STREET: 1220 Market Street, P.O. Box 2207

; CITY: Wilmington

; STATE: Delaware

; COUNTRY: U.S.A.

; ZIP: 19899

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/985,526

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/608,845

FILING DATE: 16-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: McMorriow Jr., Robert G

TELECOMMUNICATION INFORMATION:

TELEPHONE: (302) 658-9141

TELEFAX: (302) 658-5613

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-985-526-3

Query Match          20.1%; Score 28; DB 3; Length 441;
Best Local Similarity 30.8%; Pred. No. 98;
Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 308 WSEWTSCTSCGN 320

RESULT 23
US-09-092-636-9
; Sequence 9, Application US/09092636A
; Patent No. 6162641
; GENERAL INFORMATION:
; APPLICANT: Goldman, Daniel
; APPLICANT: Sapru, Mohan K.
; TITLE OF INVENTION: Neuregulin Response Element and Uses Therefor
; FILE REFERENCE: UMI-003
; CURRENT APPLICATION NUMBER: US/09/092,636A
; CURRENT FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-092-636-9

Query Match          20.1%; Score 28; DB 4; Length 468;
Best Local Similarity 30.8%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 114 WLLWATNEFSLVN 126

RESULT 24
US-09-092-636-4
; Sequence 4, Application US/09092636A
; Patent No. 6162641
; GENERAL INFORMATION:
; APPLICANT: Goldman, Daniel
; APPLICANT: Sapru, Mohan K.
; TITLE OF INVENTION: Neuregulin Response Element and Uses Therefor
; FILE REFERENCE: UMI-003
; CURRENT APPLICATION NUMBER: US/09/092,636A
; CURRENT FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-092-636-4

Query Match          20.1%; Score 28; DB 4; Length 470;
Best Local Similarity 30.8%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 114 WLHWATNEFSLVN 126

RESULT 25
```

```

US-09-295-186-18
; Sequence 18, Application US/09295186B
; Patent No. 6127137
; GENERAL INFORMATION:
; APPLICANT: Hasida, Miyoko
; APPLICANT: Tsutsumi, No. 6127137iko
; APPLICANT: Halkier, Torben
; APPLICANT: Stringer, Mary Ann
; TITLE OF INVENTION: An Acidic Phospholipase, Production, and
; TITLE OF INVENTION: Methods of Using Thereof (As Amended)
; FILE REFERENCE: 4953.204-US
; CURRENT APPLICATION NUMBER: US/09/295,186B
; CURRENT FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 1215/96
; PRIOR FILING DATE: 1996-10-31
; PRIOR APPLICATION NUMBER: PCT/DK97/00490
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Torulaspora delbrueckii
US-09-295-186-18

Query Match          20.1%; Score 28; DB 3; Length 649;
Best Local Similarity 30.8%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 156 WNNWTSVQDIVNN 168

RESULT 26
US-08-038-760-3
; Sequence 3, Application US/08038760
; Patent No. 5496731
; GENERAL INFORMATION:
; APPLICANT: Xu, Hong-Ji
; APPLICANT: Hu, Shi-Xue
; APPLICANT: Benedict, William F.
; TITLE OF INVENTION: Broad-Spectrum Tumor Suppressor Genes, Gene Products and
; TITLE OF INVENTION: Methods for Tumor Suppressor Gene Therapy.
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/038,760
; FILING DATE: 19930325
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 7409-025-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 amino acids
; TYPE: AMINO ACID
```



```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-038-760-3

Query Match      20.1%; Score 28; DB 1; Length 816;
Best Local Similarity 42.9%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      42 WXXXQTXLXNEXXL 55
      | | | | |
Db      569 WTLFQHTLQNEYEL 582

RESULT 27
US-08-470-091-3
; Sequence 3, Application US/08470091
; Patent No. 5912236
; GENERAL INFORMATION:
; APPLICANT: Xu, Hong-Ji
; APPLICANT: Hu, Shi-Xue
; APPLICANT: Benedict, William F.
; TITLE OF INVENTION: Broad-Spectrum Tumor Suppressor Genes, Gene Products and
; TITLE OF INVENTION: Methods for Tumor Suppressor Gene Therapy.
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,091
; FILING DATE: JUN-16-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/038,760
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 7409-025-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-470-091-3

Query Match      20.1%; Score 28; DB 2; Length 816;
Best Local Similarity 42.9%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      42 WXXXQTXLXNEXXL 55
      | | | | |
Db      569 WTLFQHTLQNEYEL 582

RESULT 28
US-08-204-329-1
; Sequence 1, Application US/08204329
; Patent No. 5710255
; GENERAL INFORMATION:
```

```

; APPLICANT: SHEPARD, H. M.
; APPLICANT: WEN, SHU F.
; TITLE OF INVENTION: CHARACTERIZATION OF A NOVEL ANTI-P110RB
; TITLE OF INVENTION: MONOCLONAL ANTIBODY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP
; STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,329
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05866
; FILING DATE: 14-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: RENEE A. FITTS
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 16930-0004000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 928 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-204-329-1

Query Match      20.1%; Score 28; DB 1; Length 928;
Best Local Similarity 42.9%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      42 WXXXQTXLXNEXXL 55
      | | | | |
Db      681 WTLFQHTLQNEYEL 694

RESULT 29
US-08-959-638-8
; Sequence 8, Application US/08959638
; Patent No. 5932210
; GENERAL INFORMATION:
; APPLICANT: Gregory, Richard J.
; APPLICANT: Wills, Ken N.
; APPLICANT: Maneval, Daniel C.
; TITLE OF INVENTION: Recombinant Adenoviral Vector and
; TITLE OF INVENTION: Methods of Use
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/959,638
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/328,673
FILING DATE: 25-OCT-1994
APPLICATION NUMBER: US 08/233,777
FILING DATE: 19-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/142,669
FILING DATE: 25-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-CJ 1192
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 928 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-959-638-8

Query Match 20.1%; Score 28; DB 2; Length 928;
Best Local Similarity 42.9%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNEXXL 55
| | | | |
Db 681 WTLFQHTLQNEYEL 694

RESULT 30
US-08-482-627-5
Sequence 5, Application US/08482627
Patent No. 5998134
GENERAL INFORMATION:
APPLICANT: Lee, Wen-Hwa
APPLICANT: Lee, Eva Y-H.P
TITLE OF INVENTION: Retinoblastoma Gene - Cancer Suppressor
TITLE OF INVENTION: and Regulator
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,627
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/951,947
FILING DATE: 28-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UC 1707
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 928 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-627-5
Query Match 20.1%; Score 28; DB 2; Length 928;
Best Local Similarity 42.9%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 42 WXXXQTXLXNEXXL 55
| | | | |
Db 681 WTLFQHTLQNEYEL 694

RESULT 31
US-08-801-092-4
Sequence 4, Application US/08801092
Patent No. 6074850
GENERAL INFORMATION:
APPLICANT: Antelman, Douglas
APPLICANT: Gregory, Richard J.
APPLICANT: Wils, Kenneth N.
TITLE OF INVENTION: Tissue Specific Expression of
TITLE OF INVENTION: Retinoblastoma Protein
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,092
FILING DATE: 14-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/751,517
FILING DATE: 15-NOV-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 016930-001020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 703-576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 928 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-801-092-4

Query Match 20.1%; Score 28; DB 3; Length 928;
Best Local Similarity 42.9%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 42 WXXXQTXLXNEXXL 55
| | | | |
Db 681 WTLFQHTLQNEYEL 694

RESULT 32
US-08-328-673A-8

```
; Sequence 8, Application US/08328673A
; Patent No. 6210939
; GENERAL INFORMATION:
; APPLICANT: Gregory, Richard J.
;           Wills, Ken N.
;           Maneval, Daniel C.
; TITLE OF INVENTION: Recombinant Adenoviral Vector and
;           Methods of Use
;
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,673A
; FILING DATE: 25-Oct-1994
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/142,669
; FILING DATE: 25-OCT-1993
; APPLICATION NUMBER: US 08/233,669
; FILING DATE: 26-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy S.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 016930-000920US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 8:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 928 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-08-328-673A-8

Query Match      20.1%; Score 28; DB 4; Length 928;
Best Local Similarity 42.9%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      42 WXXXQTXLXNEXXL 55
      | | | | |
Db      681 WTLFQHTLQNEVEL 694

RESULT 33
US-09-315-113-4
; Sequence 4, Application US/09315113
; Patent No. 6379927
; GENERAL INFORMATION:
; APPLICANT: Antelman, Douglas
;           Gregory, Richard J.
;           Wills, Kenneth N.
; TITLE OF INVENTION: Tissue Specific Expression of
;           Retinoblastoma Protein
;
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
```

```
;
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,113
; FILING DATE: 19-May-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/801,092
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 016930-001020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 703-576-0300
;
; INFORMATION FOR SEQ ID NO: 4:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 928 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-315-113-4

Query Match      20.1%; Score 28; DB 4; Length 928;
Best Local Similarity 42.9%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      42 WXXXQTXLXNEXXL 55
      | | | | |
Db      681 WTLFQHTLQNEVEL 694

RESULT 34
PCT-US94-10357-2
; Sequence 2, Application PC/TUS9410357
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; APPLICANT: and Canji, Inc.
; TITLE OF INVENTION: Therapeutic Use of the Retinoblastoma
; TITLE OF INVENTION: Susceptibility Gene Product
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10357
; FILING DATE: 13-SEP-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,108
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: FP-UC 1117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 928 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-10357-2

Query Match 20.1%; Score 28; DB 5; Length 928;
Best Local Similarity 42.9%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNEXXL 55
| | | | |
Db 681 WTLFQHTLQNEYEL 694

RESULT 35

PCT-US94-10357-3

; Sequence 3, Application PC/TUS9410357

; GENERAL INFORMATION:

; APPLICANT: The Regents of the University of California

; APPLICANT: and Canji, Inc.

; TITLE OF INVENTION: Therapeutic Use of the Retinoblastoma

; TITLE OF INVENTION: Susceptibility Gene Product

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell and Flores

; STREET: 4370 La Jolla Village Drive

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/10357

; FILING DATE: 13-SEP-1994

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/121,108

; FILING DATE: 13-SEP-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.

; REGISTRATION NUMBER: 31,815

; REFERENCE/DOCKET NUMBER: FP-UC 1117

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 535-9001

; TELEFAX: (619) 535-8949

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 928 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

PCT-US94-10357-3

Query Match 20.1%; Score 28; DB 5; Length 928;
Best Local Similarity 42.9%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNEXXL 55
| | | | |
Db 681 WTLFQHTLQNEYEL 694

RESULT 36

US-09-302-812-6

; Sequence 6, Application US/09302812B

; Patent No. 6333148

; GENERAL INFORMATION:

; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AM, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG)
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACT
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: NIAD 201
; CURRENT APPLICATION NUMBER: US/09/302,812B
; CURRENT FILING DATE: 1999-04-30
; EARLIER APPLICATION NUMBER: 60/083,768
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 6
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
US-09-302-812-6

Query Match 20.1%; Score 28; DB 4; Length 968;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNE 52
| | | | |
Db 528 WELIQTALLNK 538

RESULT 37

US-09-511-477-6

; Sequence 6, Application US/09511477

; Patent No. 6337202

; GENERAL INFORMATION:

; APPLICANT: JACOBSON, Myron K.

; APPLICANT: JACOBSON, Elaine L.

; APPLICANT: AME, Jean-Christophe

; APPLICANT: LIN, Winston

; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG)

; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREAC

; TITLE OF INVENTION: THEREWITH

; FILE REFERENCE: NIAD 201

; CURRENT APPLICATION NUMBER: US/09/511,477

; CURRENT FILING DATE: 2000-02-23

; PRIOR APPLICATION NUMBER: 09/302,812

; PRIOR FILING DATE: 1999-04-30

; NUMBER OF SEQ ID NOS: 38

; SEQ ID NO 6

; LENGTH: 968

; TYPE: PRT

; ORGANISM: Mus musculus

; FEATURE:

US-09-511-477-6

Query Match 20.1%; Score 28; DB 4; Length 968;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNE 52
| | | | |
Db 528 WELIQTALLNK 538

RESULT 38

US-09-511-507-6

; Sequence 6, Application US/09511507

; Patent No. 6395543

; GENERAL INFORMATION:

; APPLICANT: JACOBSON, Myron K.

; APPLICANT: JACOBSON, Elaine L.

; APPLICANT: AME, Jean-Christophe

; APPLICANT: LIN, Winston

; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG)

```

; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIV
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: NIAD 201
; CURRENT APPLICATION NUMBER: US/09/511,507
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/302,812
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 6
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
US-09-511-507-6

Query Match      20.1%; Score 28; DB 4; Length 968;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      42 WXXXQTXLXNE 52
      | ||| | |
Db      528 WELIQTALLNK 538

RESULT 39
US-09-302-812-4
; Sequence 4, Application US/09302812B
; Patent No. 6333148
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AM, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG) ENZ
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: NIAD 201
; CURRENT APPLICATION NUMBER: US/09/302,812B
; CURRENT FILING DATE: 1999-04-30
; EARLIER APPLICATION NUMBER: 60/083,768
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 4
; LENGTH: 976
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-302-812-4

Query Match      20.1%; Score 28; DB 4; Length 976;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      42 WXXXQTXLXNE 52
      | ||| | |
Db      536 WELIQTALLNK 546

RESULT 40
US-09-511-477-4
; Sequence 4, Application US/09511477
; Patent No. 6337202
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AME, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG) EN
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIV
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: NIAD 201
; CURRENT APPLICATION NUMBER: US/09/511,477
; CURRENT FILING DATE: 2000-02-23

```

```

; PRIOR APPLICATION NUMBER: 09/302,812
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 4
; LENGTH: 976
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-511-477-4

Query Match      20.1%; Score 28; DB 4; Length 976;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      42 WXXXQTXLXNE 52
      | ||| | |
Db      536 WELIQTALLNK 546

RESULT 41
US-09-511-507-4
; Sequence 4, Application US/09511507
; Patent No. 6395543
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AME, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG) ENZ
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIV
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: NIAD 201
; CURRENT APPLICATION NUMBER: US/09/511,507
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/302,812
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 4
; LENGTH: 976
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-511-507-4

Query Match      20.1%; Score 28; DB 4; Length 976;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      42 WXXXQTXLXNE 52
      | ||| | |
Db      536 WELIQTALLNK 546

RESULT 42
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and AviHu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```



```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-20

```

Query Match 20.1%; Score 28; DB 1; Length 1170;
 Best Local Similarity 30.8%; Pred. No. 2.5e+02;
 Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 385 WSEWTSCTSCGN 397

```

```

RESULT 43
US-09-342-653-7
; Sequence 7, Application US/09342653
; Patent No. 6306632
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Chromatin Associated Proteins
; FILE REFERENCE: BB-1118
; CURRENT APPLICATION NUMBER: US/09/342,653
; CURRENT FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: 60/092,841
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-342-653-7

```

Query Match 19.4%; Score 27; DB 4; Length 418;
 Best Local Similarity 36.4%; Pred. No. 1.6e+02;
 Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 39 WXXWXXXQTXL 49
| | | | |
Db 66 WLWLGLQAAL 76

```

```

RESULT 44
US-08-850-880-2
; Sequence 2, Application US/08850880
; Patent No. 5925545
; GENERAL INFORMATION:
; APPLICANT: Reznikoff, William S
; APPLICANT: Gorysin, Igor Y
; APPLICANT: Zhou, Hong
; TITLE OF INVENTION: System for In Vitro Transposition
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady

```

```

; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,880
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.94142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608/251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-850-880-2

```

Query Match 19.4%; Score 27; DB 2; Length 476;
 Best Local Similarity 36.4%; Pred. No. 1.8e+02;
 Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

QY 39 WXXWXXXQTXL 49
| | | | |
Db 450 WEGWEALQSKL 460

```

```

RESULT 45
US-08-944-916-2
; Sequence 2, Application US/08944916
; Patent No. 5948622
; GENERAL INFORMATION:
; APPLICANT: Reznikoff, William S
; APPLICANT: Goryshin, Igor Y
; APPLICANT: York, Dona L
; APPLICANT: Zhou, Hong
; TITLE OF INVENTION: System for In Vitro Transposition
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,916
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/814,877
; FILING DATE: 09-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/850,880
; FILING DATE: 02-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J

```

; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.94916
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608/251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-944-916-2

Query Match 19.4%; Score 27; DB 2; Length 476;
Best Local Similarity 36.4%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | |
Db 450 WEGWEALQSKL 460

RESULT 46

US-08-814-877-2
; Sequence 2, Application US/08814877
; Patent No. 5965443
; GENERAL INFORMATION:
; APPLICANT: Goryshin, Igor Y
; APPLICANT: Reznikoff, William S
; TITLE OF INVENTION: System for In Vitro Transposition
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,877
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.94142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608/251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-814-877-2

Query Match 19.4%; Score 27; DB 2; Length 476;
Best Local Similarity 36.4%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | |
Db 450 WEGWEALQSKL 460

RESULT 47

US-09-632-021-2

; Sequence 2, Application US/09632021
; Patent No. 6406896
; GENERAL INFORMATION:
; APPLICANT: Reznikoff, William S
; APPLICANT: Naumann, Todd A
; TITLE OF INVENTION: Transposase Enzyme and Method for Use
; FILE REFERENCE: 960296.96471
; CURRENT APPLICATION NUMBER: US/09/632,021
; CURRENT FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/146686
; PRIOR FILING DATE: 1999-08-02
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Transposon Tn5
US-09-632-021-2

Query Match 19.4%; Score 27; DB 4; Length 476;
Best Local Similarity 36.4%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | |
Db 450 WEGWEALQSKL 460

RESULT 48

US-09-272-432A-2
; Sequence 2, Application US/09272432A
; Patent No. 6437109
; GENERAL INFORMATION:
; APPLICANT: Reznikoff, William S
; APPLICANT: Goryshin, Igor Y
; APPLICANT: Zhou, Hong
; TITLE OF INVENTION: System for In Vitro Transposition
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,432A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.96277
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608/251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-272-432A-2

Query Match 19.4%; Score 27; DB 4; Length 476;
Best Local Similarity 36.4%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | |
Db 450 WEGWEALQSKL 460

RESULT 49

US-08-625-322-2
; Sequence 2, Application US/08625322
; Patent No. 5804412
; GENERAL INFORMATION:
; APPLICANT: Gill, Gordon N.
; APPLICANT: Kurten, Richard C.
; APPLICANT: Cadena, Deborah L.
; TITLE OF INVENTION: Sorting Nexins and Methods of Using Same
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,322
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1955
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 522 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-625-322-2

Query Match 19.4%; Score 27; DB 1; Length 522;
Best Local Similarity 36.4%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | |
Db 417 WQRWQDAQATL 427

RESULT 50

US-08-173-508-2
; Sequence 2, Application US/08173508
; Patent No. 5616485
; GENERAL INFORMATION:
; APPLICANT: Bartfeld, Daniel
; APPLICANT: Butler, Michael J.
; APPLICANT: Hadary, Dany
; APPLICANT: Jenish, David
; APPLICANT: Krieger, Timothy
; TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
; TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.

COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,508
FILING DATE: 23-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/125/CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 537 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-173-508-2

Query Match 19.4%; Score 27; DB 1; Length 537;
Best Local Similarity 28.6%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 412 WRTWDRDNTLRLHRD 425

Search completed: July 18, 2003, 15:51:45
Job time : 18 secs

THIS PAGE BLANK (USPTO)

Db 1 MARTVFCEYLKKEAGLDQLYPGELGKRIFDSVSKQAWGEWIKKQTMLVNEKKL 55

RESULT 2

Y553_BUCAI STANDARD; PRT; 77 AA.

AC P57618;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein BU553.

GN BU553.

OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum

OS symbiotic bacterium).

OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.

OX NCBI_TaxID=118099;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Tokyo 1998;

RX MEDLINE=20445173; PubMed=10993077;

RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;

RT "Genome sequence of the endocellular bacterial symbiont of aphids

RT Buchnera sp. APS.";

RL Nature 407:81-86(2000).

CC -!- SIMILARITY: STRONG, TO E.COLI YGGX AND H.INFLUENZAE HI0760.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AP001119; BAB13245.1; ALT_INIT.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 77 AA; 9511 MW; 00B049027CF480BF CRC64;

Query Match 33.1%; Score 46; DB 1; Length 77;

Best Local Similarity 23.6%; Pred. No. 0.0013;

Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Oy 1 MXXRXXCXXXXXXXPPXXXXXXGXXXXXXXWXXXQTXLXNEXXL 55

Db 1 MNRIFFTKKSEGDQFQSGKGLKKIYDQISKRAWEKWKIEQTILINEENL 55

RESULT 3

YGGX_ECOLI STANDARD; PRT; 90 AA.

AC P52065;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Protein yggX.

GN YGGX OR B2962 OR Z4307 OR ECS3838.

OS Escherichia coli, and

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI_TaxID=562, 83334;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / EDL933 / ATCC 700927;

RX MEDLINE=21074935; PubMed=11206551;

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,

RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,

RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,

RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

RA Welch R.A., Blattner F.R.;

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";

RL Nature 409:529-533(2001).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / RIMD 0509952;

RX MEDLINE=21156231; PubMed=11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,

RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,

RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,

RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;

RT "Complete genome sequence of enterohaemorrhagic Escherichia coli

RT O157:H7 and genomic comparison with a laboratory strain K-12.";

RL DNA Res. 8:11-22(2001).

RN [4]

RP PARTIAL SEQUENCE OF 1-12.

RC STRAIN=K12 / EMG2;

RX MEDLINE=97443975; PubMed=9298646;

RA Link A.J., Robison K., Church G.M.;

RT "Comparing the predicted and observed properties of proteins encoded

RT in the genome of Escherichia coli K-12.";

RL Electrophoresis 18:1259-1313(1997).

RN [5]

RP SEQUENCE OF 1-10.

RC STRAIN=K12;

RX MEDLINE=99085675; PubMed=9868784;

RA Wasinger V.C., Humphrey-Smith I.;

RT "Small genes/gene-products in Escherichia coli K-12.";

RL FEMS Microbiol. Lett. 169:375-382(1998).

RN [6]

RP IDENTIFICATION BY MASS SPECTROMETRY.

RX MEDLINE=99420866; PubMed=10493123;

RA Fountoulakis M., Takacs M.-F., Berndt P., Langen H., Takacs B.;

RT "Enrichment of low abundance proteins of Escherichia coli by

RT hydroxyapatite chromatography.";

RL Electrophoresis 20:2181-2195(1999).

CC -!- SIMILARITY: STRONG, TO H.INFLUENZAE HI0760.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U28377; AAA69129.1; -.

DR EMBL; AE000378; AAC75999.1; -.

DR EMBL; AE005526; AAG58093.1; -.

DR EMBL; AP002563; BAB37261.1; -.

DR SWISS-2DPAGE; P52065; COLI.

DR EcoGene; EGI2984; yggX.

KW Complete proteome.

FT INIT_MET 0

SQ SEQUENCE 90 AA; 10821 MW; D7C66C2A35E62402 CRC64;

Query Match 31.7%; Score 44; DB 1; Length 90;

Best Local Similarity 30.3%; Pred. No. 0.004;

Matches 10; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Oy 23 PXXGXXXXXXXWXXXQTXLXNEXXL 55

Db 22 PGELGKRIYNEISKEAWAQWQHQTMLINEKKL 54

RESULT 4
PHBC_ALCEU STANDARD; PRT; 589 AA.
AC P23608;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Poly-beta-hydroxybutyrate polymerase (EC 2.3.1.-) (Poly(3-hydroxybutyrate) polymerase) (PHB polymerase) (PHA synthase)
DE hydroxalkanoate) polymerase) (PHA polymerase) (PHA synthase)
DE (Polyhydroxyalkanoic acid synthase).
GN PHBC.
OS Alkaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H16 / DSM 428 / ATCC 17699;
RX MEDLINE=89359357; PubMed=2670936;
RA Peoples O.P., Sinskey A.J.;
RT "Poly-beta-hydroxybutyrate (PHB) biosynthesis in Alkaligenes eutrophus H16. Identification and characterization of the PHB polymerase gene (phbC).";
RL J. Biol. Chem. 264:15298-15303(1989).
RN [2]
RP SEQUENCE OF 1-219 FROM N.A.
RX MEDLINE=91100279; PubMed=19871116;
RA Schubert P., Krueger N., Steinbuechel A.;
RT "Molecular analysis of the Alkaligenes eutrophus poly(3-hydroxybutyrate) biosynthetic operon: identification of the N terminus of poly(3-hydroxybutyrate) synthase and identification of the promoter.";
RL J. Bacteriol. 173:168-175(1991).
CC -1- FUNCTION: POLYMERIZES D(-)-3-HYDROXYBUTYRYL-COA TO CREATE PHB WHICH CONSISTS OF THOUSANDS OF HYDROXYBUTYRATE MOLECULES LINKED END TO END. PHB SERVES AS AN INTRACELLULAR ENERGY RESERVE MATERIAL WHEN CELLS GROW UNDER CONDITIONS OF NUTRIENT LIMITATION.
CC -1- PATHWAY: Poly-beta-hydroxybutyrate biosynthesis; third step.
CC -1- SUBUNIT: MONOMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- BIOTECHNOLOGY: Forms a biodegradable plastic that is degraded naturally and completely by bacteria into carbon dioxide and water. Utilized in the medical industry. Plates made from PHA-based plastics can be left in place to help heal fractured bones. After the bone has healed, the plastic slowly breaks down in the body. Utilized by Imperial Chemical Industries (ICI) to produce a PHB-PHV (poly-B-valerate) copolymer sold under the trade name 'Biopol'. Biopol is used as packaging material. The PHB-PHV copolymer consists of approximately 20% PHV and 80% PHB. It can be synthesized by incorporating glucose and valeric acid into the medium. PHB-PHV is stronger and more flexible than regular PHB. Under industrial conditions, 80% or higher of the cell dry weight of A.eutrophus usually consists of the PHB-PHV copolymer.
CC -1- SIMILARITY: BELONGS TO THE PHA/PHB SYNTHASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; J05003; AAA21975.1; -.
DR EMBL; M64341; AAA21979.1; -.
DR PIR; A34341; A34341.
DR InterPro; IPR000073; Abhydrolase.
DR Pfam; PF00561; abhydrolase; 1.
KW PHB biosynthesis; transferase; Acyltransferase.
FT ACT_SITE 319 319 POTENTIAL.
SQ SEQUENCE 589 AA; 64316 MW; A822F35CF70D8B68 CRC64;

Query Match 24.5%; Score 34; DB 1; Length 589;
Best Local Similarity 42.9%; Pred. No. 3.5;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 436 WYCWLRLRHTYLQNE 449

RESULT 5
YEDQ_ECO57 STANDARD; PRT; 564 AA.
AC Q8XB92;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yedQ.
GN YEDQ OR Z3047 OR ECS2694.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7." Nature 409:529-533(2001).
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasakiwa C., Ogasawara N., Yasunaga T., Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12." DNA Res. 8:11-22(2001).
RL DNA Res. 8:11-22(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE YAIC / YFIN (E.COLI), YHCK (B.SUBTILIS) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AE005417; AAG56970.1; -.
CC EMBL; AP002559; BAB36117.1; -.
CC InterPro; IPR000160; GGDEF.
CC Pfam; PF00990; GGDEF; 1.
CC SMART; SM00267; DUF1; 1.
CC TIGRFAMS; TIGR00254; GGDEF; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 20 40 POTENTIAL.
FT TRANSMEM 360 380 POTENTIAL.
SQ SEQUENCE 564 AA; 64287 MW; BEBC2286ADBAECB0 CRC64;

Query Match 23.7%; Score 33; DB 1; Length 564;
Best Local Similarity 35.3%; Pred. No. 5.6;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXLXNEXXL 55

Db 396 WQAWHDTLTRLYNRGAL 412

RESULT 6

YEDQ_ECOLI
ID YEDQ_ECOLI STANDARD; PRT; 564 AA.
AC P76330; P94746;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yedQ.
GN YEDQ OR B1956.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
Nishimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
Sivasubram S., Tagami H., Takeda J., Takemoto K., Wada C.,
Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA Sequence of the Escherichia coli K-12 Genome
Corresponding to the 40.1-50.0 min Region on the Linkage Map.";
RL DNA Res. 3:379-392(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE YAIC / YFIN (E.COLI), YHCK (B.SUBTILIS)
FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE000287; AAC75022.1; ALT_INIT.
DR EMBL; D90835; BAA15784.1; -.
DR EcoGene; EG14040; yedQ.
DR InterPro; IPR000160; GGDEF.
DR Pfam; PF00990; GGDEF; 1.
DR SMART; SM00267; DUF1; 1.
DR TIGRFAMS; TIGR00254; GGDEF; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 20 40 POTENTIAL.
FT TRANSMEM 360 380 POTENTIAL.
SQ SEQUENCE 564 AA; 64283 MW; 05FB02C1BE2A8938 CRC64;

Query Match 23.7%; Score 33; DB 1; Length 564;
Best Local Similarity 35.3%; Pred. No. 5.6;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQFXLXNEXXL 55
Db 396 WQAWHDTLTRLYNRGAL 412

RESULT 7

YAMB_SCHPO
ID YAMB_SCHPO STANDARD; PRT; 3655 AA.
AC Q10064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ClF5.11c in chromosome I.
GN SPAC1F5.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: STRONG, TO YEAST YHR099W.
CC -!- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on i
CC use by non-profit institutions as long as its content is in no w
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z68136; CAA92239.1; -.
DR InterPro; IPR003151; FAT.
DR InterPro; IPR003152; FATC.
DR InterPro; IPR000403; PI3_PI4_kinase.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00454; PI3_PI4_kinase; 1.
DR Pfam; PF02259; FAT; 1.
DR Pfam; PF02260; FATC; 1.
DR SMART; SM00146; PI3K; 1.
DR PROSITE; PS50290; PI3_4_KINASE_3; 1.
KW Hypothetical protein; Transferase; Kinase.
FT DOMAIN 3324 3655 PI3K/PI4K (BY SIMILARITY).
SQ SEQUENCE 3655 AA; 420774 MW; 50475E3F3FC2124A CRC64;

Query Match 23.0%; Score 32; DB 1; Length 3655;
Best Local Similarity 54.5%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNE 52
| | | | |
Db 1553 WLFFQTILTNE 1563

RESULT 8

LEF7_NPVOP STANDARD; PRT; 211 AA.
ID LEF7_NPVOP
AC O10362;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Late expression factor 7.
GN LEF-7.
OS Orgyia pseudotsugata multicapsid polyhedrosis virus (OpMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=164623;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97271300; PubMed=9126251;
RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA Rohrmann G.F.;
RT "The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear
polyhedrosis virus genome.";
RL Virology 229:381-399(1997).
CC -!- FUNCTION: INVOLVED IN LATE/VERY LATE GENE ACTIVATION (BY
SIMILARITY).
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; U75930; AAC59122.1; .
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
KW Transcription regulation.
SQ SEQUENCE 211 AA; 24139 MW; 9D5731720657E8AB CRC64;

Query Match 22.3%; Score 31; DB 1; Length 211;
Best Local Similarity 35.7%; Pred. No. 6.5;
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXXQTXLXNE 52
| | | | |
Db 133 WNWGLTRTLIHE 146

RESULT 9

YEDQ_SALTI STANDARD; PRT; 567 AA.
ID YEDQ_SALTI
AC Q8Z5R0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yedQ.
GN YEDQ OR STY2194.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,

RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
RT Nature 413:848-852(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE YAIC / YFIN (E.COLI), YHCK (B.SUBTILIS)
FAMILY.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; AL627272; CAD05734.1; .
DR InterPro; IPR000160; GGDEF.
DR Pfam; PF00990; GGDEF; 1.
DR SMART; SM00267; DUF1; 1.
DR TIGRFAMS; TIGR00254; GGDEF; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 20 40 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
SQ SEQUENCE 567 AA; 65099 MW; 9E1AE3768ADADF6D CRC64;

Query Match 22.3%; Score 31; DB 1; Length 567;
Best Local Similarity 35.3%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXXQTXLXNEXXL 55
| | | | |
Db 393 WQAWHDLTRLNRGAL 409

RESULT 10

YEDQ_SALTY STANDARD; PRT; 570 AA.
ID YEDQ_SALTY
AC Q8ZNT5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yedQ.
GN YEDQ OR STM1987.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RT Nature 413:852-856(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE YAIC / YFIN (E.COLI), YHCK (B.SUBTILIS)
FAMILY.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; AF008788; AAL20897.1; -.
DR StyGene; SG????; yedQ.
DR InterPro; IPR000160; GGDEF.
DR Pfam; PF00990; GGDEF; 1.
DR SMART; SM00267; DUF1; 1.
DR TIGRFAMS; TIGR00254; GGDEF; 1.
KW Hypothetical protein; Transmembrane; Complete proteome:
FT TRANSMEM 20 40 POTENTIAL.
FT TRANSMEM 360 380 POTENTIAL.
SQ SEQUENCE 570 AA; 65429 MW; BF8FDF0CF8949925 CRC64;

Query Match 22.3%; Score 31; DB 1; Length 570;
Best Local Similarity 35.3%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
Db 396 WQAWHDPLTRLNRYNGAL 412

RESULT 11
SLT_HAEIN STANDARD; PRT; 593 AA.
AC P44888;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative soluble lytic murein transglycosylase precursor (EC 3.2.1.-).
GN SLT OR HI0829.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: MUREIN-DEGRADING ENZYME. MAY PLAY A ROLE IN RECYCLING
CC OF MUROPEPTIDES DURING CELL ELONGATION AND/OR CELL DIVISION
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Cleavage of the beta-1,4-glycosidic bond
CC between N-acetylmuramic acid and N-acetylglucosamine residues,
CC thereby conserving the energy in a newly synthesized
CC 1,6-anhydrobond in the muramic acid residue.
CC -!- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC -!- SIMILARITY: BELONGS TO THE SLT FAMILY OF TRANSGLYCOSYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32765; AAC22487.1; -.
DR HSSP; P03810; IQSA.
DR TIGR; HI0829; -.
DR InterPro; IPR000189; SLT_domain.

DR Pfam; PF01464; SLT; 1.
DR PROSITE; PS00922; TRANSGLYCOSYLASE; 1.
KW Cell wall; Hydrolase; Glycosidase; Periplasmic; Signal;
KW Complete proteome.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 593 PUTATIVE SOLUBLE LYTIC MUREIN
FT DOMAIN 440 529 TRANSGLYCOSYLASE.
FT ACT_SITE 453 453 SLT-TYPE DOMAIN.
FT SEQUENCE 593 AA; 68691 MW; 774DDD3D38217CEE CRC64;

Query Match 22.3%; Score 31; DB 1; Length 593;
Best Local Similarity 35.3%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
Db 300 WRYWEAKQDILKNTKKL 316

RESULT 12
YQJK_ECOLI STANDARD; PRT; 99 AA.
AC Q47710;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yqjK.
GN YQJK OR B3100.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18997; AAA57904.1; -.
DR EMBL; AE000392; AAC76135.1; -.
DR EcoGene; EG14314; yqjK.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 99 AA; 11811 MW; DAD56718FFAF4086 CRC64;

Query Match 21.6%; Score 30; DB 1; Length 99;
Best Local Similarity 28.6%; Pred. No. 5.5;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
Db 82 WSAWRLVKTTLKQ 95

RESULT 13
YB94_METTH STANDARD; PRT; 331 AA.
ID YB94_METTH
AC O27262;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)


```
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MTH1194.
GN MTH1194.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell J., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000887; AAB85683.1; -
DR InterPro; IPR00286; His_deacetylase.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
KW Hypothetical protein; Hydrolase; Complete proteome.
SQ SEQUENCE 331 AA; 36722 MW; 489054F32965EDCF CRC64;

Query Match 21.6%; Score 30; DB 1; Length 331;
Best Local Similarity 45.5%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | |
Db 263 WIGWFIHQTGL 273

RESULT 14
YE87_SCHPO
ID YE87_SCHPO STANDARD; PRT; 418 AA.
AC O14303;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C9G1.07 in chromosome I.
GN SPAC9G1.07.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourous J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
```

```
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z98763; CAB11491.1; -
KW Hypothetical protein.
SQ SEQUENCE 418 AA; 46357 MW; C01D2DE4466C03E1 CRC64;

Query Match 21.6%; Score 30; DB 1; Length 418;
Best Local Similarity 29.4%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLNEXXL 55
| | | | |
Db 319 WGSWVSSQDTSKNSSNL 335

RESULT 15
ETS2_CHICK
ID ETS2_CHICK STANDARD; PRT; 479 AA.
AC P10157;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-ETS-2 protein.
GN ETS2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=88283637; PubMed=3293999;
RA Boulukos K.E., Pognonec P., Begue A., Galibert F., Gesquiere J.C.,
RA Stehelin D., Ghysdael J.;
RT "Identification in chickens of an evolutionarily conserved cellular
RT ets-2 gene (c-ets-2) encoding nuclear proteins related to the
RT products of the c-ets proto-oncogene.";
RL EMBO J. 7:697-705(1988).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -!- SIMILARITY: CONTAINS 1 POINTED (PNT) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X07202; CAA30178.1; -.
DR PIR; S00386; TVCHE2.
DR HSSP; P14921; 2STT.
DR TRANSFAC; T00116; -.
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ET5.
DR InterPro; IPR003118; SAM_PNT.
DR Pfam; PF00178; Ets; 1.
DR Pfam; PF02198; SAM_PNT; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR SMART; SM00251; SAM_PNT; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS0061; ETS_DOMAIN_3; 1.
KW Proto-oncogene; DNA-binding; Nuclear protein.
FT DOMAIN 90 173 POINTED.
FT DNA_BIND 373 453 ETS-DOMAIN.
SQ SEQUENCE 479 AA; 54540 MW; 91BCD5206972E867 CRC64;

Query Match 21.6%; Score 30; DB 1; Length 479;
Best Local Similarity 30.8%; Pred. No. 22;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
Db 117 WLAWATNEFSLAN 129

RESULT 16
SULA_ECOLI
ID SULA_ECOLI STANDARD; PRT; 169 AA.
AC P08846; P03840; P71224;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell division inhibitor.
GN SULA OR SPIA OR B0958 OR Z1308 OR ECS1042.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248093; PubMed=3297925;
RA Freudl R., Braun G., Honore N., Cole S.T.;
RT "Evolution of the enterobacterial sula gene: a component of the SOS
RL system encoding an inhibitor of cell division.";
RL Gene 52:31-40(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=81053729; PubMed=6253901;
RA Beck E., Bremer E.;
RT "Nucleotide sequence of the gene ompA coding the outer membrane
RL protein II of Escherichia coli K-12.";
RL Nucleic Acids Res. 8:3011-3024(1980).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]

RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobé T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [7]
RP SEQUENCE OF 1-9 FROM N.A.
RX MEDLINE=83244178; PubMed=6306396;
RA Cole S.T.;
RT "Characterisation of the promoter for the LexA regulated sula gene of
RT Escherichia coli.";
RL Mol. Gen. Genet. 189:400-404(1983).
CC -!- FUNCTION: SULA IS A COMPONENT OF THE SOS SYSTEM AND AN INHIBITOR
CC OF CELL DIVISION. ACCUMULATION OF SULA CAUSES RAPID CESSATION OF
CC CELL DIVISION AND THE APPEARANCE OF LONG, NON-SEPTATE FILAMENTS.
CC THE EXPRESSION OF SULA IS REPPRESSED BY LEXA PROTEIN. FTSZ SEEMS
CC TO BE THE TARGET OF SULA.
CC -!- SUBCELLULAR LOCATION: Inner membrane.
CC -!- MISCELLANEOUS: THE N-TERMINUS OF SULA PROTEIN MIGHT BE INVOLVED IN
CC RECOGNIZING THE CELL DIVISION APPARATUS.
CC -!- SIMILARITY: THE C-TERMINUS SHOWS SIMILARITY TO THE N PROTEIN OF
CC PHAGE LAMBDA.
CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 145
CC ONWARD DUE TO A FRAMESHIFT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00307; CAA23587.1; ALT_FRAME.
DR EMBL; V00358; CAA23654.1; -.
DR EMBL; AE000198; AAC74044.1; -.
DR EMBL; D90733; BAA35716.1; -.
DR EMBL; D90734; BAA35723.1; -.
DR EMBL; AE005286; AAG55444.1; -.
DR EMBL; AP002554; BAB34465.1; -.
DR PIR; A04468; QOECAL.
DR PIR; A29016; A29016.
DR EcoGene; EG10984; sula.

DR InterPro; IPR004596; Sula.
DR TIGRFAMS; TIGR00623; sula; 1.
KW Cell division; Septation; SOS response; Inner membrane;
KW Complete proteome.
FT DOMAIN 31 149 CONSERVED REGION.
FT SIMILAR 150 169 TO N PROTEIN OF PHAGE LAMBDA.
FT DOMAIN 160 169 LON PROTEIN BINDING SITE (PROBABLE).
SQ SEQUENCE 169 AA; 18801 MW; C76B4493773C77C2 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 169;
Best Local Similarity 35.7%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 63 WQLWLTPQOKLSRE 76

RESULT 17
SULA_ENTAE STANDARD; PRT; 169 AA.
AC P08848;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell division inhibitor.
GN SULA.
OS Enterobacter aerogenes (Aerobacter aerogenes).
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Enterobacter.
OX NCBI_TaxID=548;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248093; PubMed=3297925;
RA Freudl R., Braun G., Honore N., Cole S.T.;
RT "Evolution of the enterobacterial sula gene: a component of the SOS
RT system encoding an inhibitor of cell division.";
RL Gene 52:31-40(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
CC -!- FUNCTION: SULA IS A COMPONENT OF THE SOS SYSTEM AND AN INHIBITOR
CC OF CELL DIVISION. ACCUMULATION OF SULA CAUSES RAPID CESSATION OF
CC CELL DIVISION AND THE APPEARANCE OF LONG, NON-SEPTATE FILAMENTS.
CC THE EXPRESSION OF SULA IS REPPRESSED BY LEXA PROTEIN. FTSZ SEEMS
CC TO BE THE TARGET OF SULA.
CC -!- SUBCELLULAR LOCATION: Inner membrane.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16467; AAA24810.1; -
CC PIR; C29016; C29016.
CC InterPro; IPR004596; Sula.
CC TIGRFAMS; TIGR00623; sula; 1.
KW Cell division; Septation; SOS response; Inner membrane.
FT DOMAIN 31 149 CONSERVED REGION.
FT SIMILAR 150 169 TO N PROTEIN OF PHAGE LAMBDA.
FT DOMAIN 160 169 LON PROTEIN BINDING SITE (PROBABLE).
SQ SEQUENCE 169 AA; 19180 MW; 3099CCD8AC568A2F CRC64;

Query Match 20.9%; Score 29; DB 1; Length 169;
Best Local Similarity 35.7%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |

Db 63 WQLWLTPQOKLSKE 76

RESULT 18
SULA_SALTY STANDARD; PRT; 169 AA.
AC P08847;
DT 01-NOV-1988 (Rel. 09, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell division inhibitor.
GN SULA OR STM1071 OR STY1092.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium;
RX MEDLINE=87248093; PubMed=3297925;
RA Freudl R., Braun G., Honore N., Cole S.T.;
RT "Evolution of the enterobacterial sula gene: a component of the SOS
RT system encoding an inhibitor of cell division.";
RL Gene 52:31-40(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
CC -!- FUNCTION: SULA IS A COMPONENT OF THE SOS SYSTEM AND AN INHIBITOR
CC OF CELL DIVISION. ACCUMULATION OF SULA CAUSES RAPID CESSATION OF
CC CELL DIVISION AND THE APPEARANCE OF LONG, NON-SEPTATE FILAMENTS.
CC THE EXPRESSION OF SULA IS REPPRESSED BY LEXA PROTEIN. FTSZ SEEMS
CC TO BE THE TARGET OF SULA.
CC -!- SUBCELLULAR LOCATION: Inner membrane.
CC -!- MISCELLANEOUS: THE N-TERMINUS OF SULA PROTEIN MIGHT BE INVOLVED IN
CC RECOGNIZING THE CELL DIVISION APPARATUS.
CC -!- SIMILARITY: THE C-TERMINUS SHOWS SIMILARITY TO THE N PROTEIN OF
CC PHAGE LAMBDA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16324; AAA27230.1; -
CC EMBL; AE008746; AAL20004.1; -
DR

DR EMBL; AL627269; CAD08197.1; -.
DR PIR; B29016; B29016.
DR StyGene; SG10386; sula.
DR InterPro; IPR004596; Sula.
DR TIGRFAMS; TIGR00623; sula; 1.
KW Cell division; Septation; SOS response; Inner membrane;
KW Complete proteome.
FT DOMAIN 31 149 CONSERVED REGION.
FT SIMILAR 150 169 TO N PROTEIN OF PHAGE LAMBDA.
FT DOMAIN 160 169 LON PROTEIN BINDING SITE (PROBABLE).
FT CONFLICT 102 102 S -> T (IN REF. 1).
FT CONFLICT 108 108 R -> A (IN REF. 1).
FT CONFLICT 151 151 H -> L (IN REF. 1).
SQ SEQUENCE 169 AA; 19013 MW; 3848A73595E5D176 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 169;
Best Local Similarity 35.7%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 63 WQLWLTPQOKLSRE 76

RESULT 19
VP10_RDVA STANDARD; PRT; 353 AA.
AC Q85447;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nonstructural protein Pns10.
GN S10.
OS Rice dwarf virus (isolate Akita) (RDV).
OC Viruses; dsRNA viruses; Reoviridae; Fijivirus.
OX NCBI_TaxID=142803;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93277380; PubMed=8503785;
RA Suzuki N.;
RT "In vitro translation of rice dwarf phytovereovirus genome segments S4 to S10.";
RL Arch. Virol. 130:201-208(1993).
CC -!- SIMILARITY: TO WOUND TUMOR VIRUS S10.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; D10221; BAA01073.1; -.
KW Nonstructural protein.
SQ SEQUENCE 353 AA; 39196 MW; AF6C2AE868186881 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 353;
Best Local Similarity 23.5%; Pred. No. 28;
Matches 4; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 279 WKGWASTYAYMFNQEQ 295

RESULT 20
VP10_RDVF STANDARD; PRT; 353 AA.
AC Q85434; Q85441;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Nonstructural protein Pns10.
GN S10.
OS Rice dwarf virus (isolate Fujian) (RDV).
OC Viruses; dsRNA viruses; Reoviridae; Fijivirus.
OX NCBI_TaxID=142804;
RN [1]
RP SEQUENCE FROM N.A.
RA Chu R., Zhang X., Pan N., Chen Z.;
RT "Molecular cloning and sequencing of rice dwarf virus segment 10.";
RL Chin. J. Bot. 35:115-120(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Chu R., Zhang X., Pan N., Chen Z.;
RT "The cDNA cloning and nucleotide sequence of the gene encoding
RT nonstructural protein of rice dwarf virus genome segment 10.";
RL Acta Bot. Sin. 35:115-120(1993).
CC -!- SIMILARITY: TO WOUND TUMOR VIRUS S10.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U36567; AAA88766.1; -.
DR EMBL; U25671; AAA70094.1; -.
KW Nonstructural protein.
FT CONFLICT 159 159 T -> I (IN REF. 1).
FT CONFLICT 320 320 D -> H (IN REF. 1).
FT CONFLICT 328 328 S -> Y (IN REF. 1).
FT CONFLICT 335 335 A -> V (IN REF. 1).
SQ SEQUENCE 353 AA; 39247 MW; 9F6D51F539758413 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 353;
Best Local Similarity 23.5%; Pred. No. 28;
Matches 4; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | | | |
Db 279 WRGWASTYAYMFNQEQ 295

RESULT 21
VP10_RDVO STANDARD; PRT; 353 AA.
AC P16594;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nonstructural protein Pns10.
GN S10.
OS Rice dwarf virus (isolate O) (RDV).
OC Viruses; dsRNA viruses; Reoviridae; Fijivirus.
OX NCBI_TaxID=142805;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88089533; PubMed=3335833;
RA Omura T., Minobe Y., Tsuchizaki T.;
RT "Nucleotide sequence of segment S10 of the rice dwarf virus genome.";
RL J. Gen. Virol. 69:227-231(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Uyeda I., Matsumura T., Sano T., Ohshima K., Shikata E.;
RT "Nucleotide sequence of rice dwarf virus genome segment 10.";
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 63:227-230(1987).
CC -!- SIMILARITY: TO WOUND TUMOR VIRUS S10.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D00241; BAA00171.1; -.
DR EMBL; M35118; AAA47232.1; -.
KW Nonstructural protein.
FT CONFLICT 17 C -> S (IN REF. 2).
SQ SEQUENCE 353 AA; 39226 MW; 8C7EC0CB559C88D0 CRC64;

Query Match . 20.9%; Score 29; DB 1; Length 353;
Best Local Similarity 23.5%; Pred. No. 28;
Matches 4; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | : | |
Db 279 WRGWASTYAYMFNQEQ 295

RESULT 22

VSGP_EBOEC STANDARD; PRT; 364 AA.
AC P87670;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Small/secreted glycoprotein precursor (SGP).
GN GP.
OS Ebola virus (strain Eckron-76) (Ebo).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;
OC Ebola-like viruses.
OX NCBI_TaxID-129000;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=97329084; PubMed-91855597;
RA Volchkov V., Volchkova V., Eckel C., Klenk H.D., Bouloy M.,
RA Leguenno B., Feldmann H.;
RT "Emergence of subtype zaire Ebola virus in Gabon.";
RL Virology 232:139-144(1997).
CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED WHEN NO RNA-EDITING
CC -!- TAKES PLACES DURING TRANSCRIPTION.
CC -!- SIMILARITY: BELONGS TO THE FILOVIRUSES GLYCOPROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U81161; AAC57993.1; -.
DR InterPro; IPR002561; Filo_glycop.
DR Pfam; PF01611; Filo_glycop; 1.
KW Glycoprotein; Signal.

FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 364 SMALL/SECRETED GLYCOPROTEIN.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 364 AA; 41202 MW; 5D034F8DA5EE2695 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 364;
Best Local Similarity 36.4%; Pred. No. 29;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | : | |
Db 288 WAFWETKKTSL 298

RESULT 23

VSGP_EBOG4 STANDARD; PRT; 364 AA.
AC O11458;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Small/secreted glycoprotein precursor (SGP).
GN GP.
OS Ebola virus (strain Gabon-94) (Ebo).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;
OC Ebola-like viruses.
OX NCBI_TaxID-128947;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97329084; PubMed-91855597;
RA Volchkov V., Volchkova V., Eckel C., Klenk H.D., Bouloy M.,
RA Leguenno B., Feldmann H.;
RT "Emergence of subtype zaire Ebola virus in Gabon.";
RL Virology 232:139-144(1997).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE FILOVIRUSES GLYCOPROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U77384; AAC57990.1; -.
DR InterPro; IPR002561; Filo_glycop.
DR Pfam; PF01611; Filo_glycop; 1.
KW Glycoprotein; Signal.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 364 SMALL/SECRETED GLYCOPROTEIN.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 364 AA; 41218 MW; 7B46B128FA3E82A5 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 364;
Best Local Similarity 36.4%; Pred. No. 29;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | : | |
Db 288 WAFWETKKTSL 298

RESULT 24

VSGP_EBOZM STANDARD; PRT; 364 AA.
AC Q66819; O12717; O12421; Q66801; Q9YMG3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Small/secreted glycoprotein precursor (SGP).
GN GP.
OS Ebola virus (strain Zaire Mayinga) (Ebo), and
OS Ebola virus (strain Zaire-95) (Ebo).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae;
OC Ebola-like viruses.
OX NCBI_TaxID-128952, 128951;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Zaire Mayinga;

RX MEDLINE=94055391; PubMed=8237108;
RA Sanchez A., Kiley M.P., Holloway B.P., Auperin D.D.;
RT "Sequence analysis of the Ebola virus genome: organization, genetic
RT elements, and comparison with the genome of Marburg virus.";
RL Virus Res. 29:215-240(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Zaire Mayinga;
RA Volchkov V.E.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Zaire-95;
RX MEDLINE=96195018; PubMed=8622982;
RA Sanchez A., Trappier S.G., Mahy B.W., Peters C.J., Nichol S.T.;
RT "The virion glycoproteins of Ebola viruses are encoded in two reading
RT frames and are expressed through transcriptional editing.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:3602-3607(1996).
RN [4]
RP PROCESSING.
RX MEDLINE=20072964; PubMed=10603327;
RA Volchkova V.A., Klenk H.D., Volchkov V.E.;
RT "Delta-peptide is the carboxy-terminal cleavage fragment of the
RT nonstructural small glycoprotein sgp of Ebola virus.";
RL Virology 265:164-171(1999).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: N- AND O-GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO THE FILOVIRUSES GLYCOPROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23187; AAC54886.1; -.
DR EMBL; U31033; AAA96745.1; -.
DR EMBL; AF086833; AAD14584.1; -.
DR EMBL; U28077; AAB37094.1; -.
DR InterPro; IPR002561; Filo_glycop.
DR Pfam; PF01611; Filo_glycop; 1.
KW Glycoprotein; Signal.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 364 SMALL/SECRETED GLYCOPROTEIN.
FT CHAIN 325 364 DELTA PEPTIDE.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 364 AA; 41175 MW; 67376A454CE5F362 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 364;
Best Local Similarity 36.4%; Pred. No. 29;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | :| |
Db 288 WAFWETKKTSL 298

RESULT 25
ENO_SYNY3
ID ENO_SYNY3 STANDARD; PRT; 432 AA.
AC P77972;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase) (2-phospho-D-
DE glycerate hydro-lyase).

GN ENO OR SLR0752.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=11148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = phosphoenolpyruvate +
CC H(2)O.
CC -!- COFACTOR: MAGNESIUM IS REQUIRED FOR CATALYSIS AND FOR STABILIZING
CC THE DIMER (BY SIMILARITY).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE ENOLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90917; BAA18749.1; -.
DR HSSP; P56252; 1PDZ.
DR InterPro; IPR000941; Enolase.
DR Pfam; PF00113; enolase; 1.
DR PRINTS; PR00148; ENOLASE.
DR ProDom; PD000902; Enolase; 1.
DR TIGRFAMS; TIGR01060; eno; 1.
DR PROSITE; PS00164; ENOLASE; 1.
KW Lyase; Glycolysis; Magnesium; Complete proteome.
FT ACT_SITE 160 160 BY SIMILARITY.
FT METAL 248 248 MAGNESIUM (BY SIMILARITY).
FT METAL 289 289 MAGNESIUM (BY SIMILARITY).
FT METAL 316 316 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 432 AA; 46528 MW; C3F347A8F0BB1B0F CRC64;

Query Match 20.9%; Score 29; DB 1; Length 432;
Best Local Similarity 29.4%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps

QY 39 WXXWXXXQTXLXNEXXL 55
| | :| |
Db 297 WESWKGLTTSLSGKTQL 313

RESULT 26
PLB1_YEAST
ID PLB1_YEAST STANDARD; PRT; 664 AA.
AC P39105;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lysophospholipase 1 precursor (EC 3.1.1.5) (Phospholipase B 1).
GN PLB1 OR YMR008C OR YMR270.10C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OX Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;

```
RX MEDLINE-94327513; PubMed=8051052;
RA Lee K.S., Patton J.L., Fido M., Hines L.K., Kohlwein S.P.,
RA Paltauf F., Henry S.A., Levin D.E.;
RT "The Saccharomyces cerevisiae PLB1 gene encodes a protein required
RT for lysophospholipase and phospholipase B activity.";
RL J. Biol. Chem. 269:19725-19730(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5288c / AB972;
RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE RELEASE OF FATTY ACIDS FROM
CC LYSOPHOSPHOLIPIDS.
CC -!- CATALYTIC ACTIVITY: 2-lysophosphatidylcholine + H(2)O =
CC glycerophosphocholine + a fatty acid anion.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: TO OTHER FUNGAL LYSOPHOSPHOLIPASES AND TO YEAST
CC SPOI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L23089; AAA61611.1; -.
DR EMBL; 248613; CAA88523.1; -.
DR PIR; A53647; A53647.
DR SGD; S0004610; PLB1.
DR InterPro; IPR002642; PLAC.
DR Pfam; PF01735; PLA2_B; 1.
DR SMART; SM00022; PLAC; 1.
DR Lipid degradation; Hydrolase; Glycoprotein; Signal.
KW SIGNAL 1 22
FT CHAIN 23 664 LYSOPHOSPHOLIPASE 1.
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 565 565 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 582 582 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 641 641 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 32 32 A -> S (IN REF. 1).
FT CONFLICT 494 494 E -> D (IN REF. 1).
SQ SEQUENCE 664 AA; 71667 MW; E05A585E7AB73F34 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 664;
Best Local Similarity 30.8%; Pred. No. 50;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 158 WNNWTSVQAIVDN 170

RESULT 27
RRPO_CARMV
ID RRPO_CARMV STANDARD; PRT; 867 AA.
```

```
AC P04518;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable RNA-directed RNA polymerase (EC 2.7.7.48).
OS Carnation mottle virus (Carmv).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
OC Carmovirus.
OC NCBI_TaxID=11986;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=86041863; PubMed=3840587;
RX Guillely H., Carrington J.C., Balazs E., Jonard G., Richards K.,
RA Morris T.J.;
RA "Nucleotide sequence and genome organization of carnation mottle
RT virus RNA.";
RT Nucleic Acids Res. 13:6663-6677(1985).
RL -!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC {RNA}(N).
CC -!- MISCELLANEOUS: READTHROUGH OF TWO TERMINATORS MAY OCCUR: UAG
CC BETWEEN CODONS AAA FOR 245-LYS AND GGG FOR 246-GLY, AND UAG
CC BETWEEN CODONS CAG FOR 762-GLN AND UUG FOR 763-LEU.
CC -!- SIMILARITY: THE RNA POLYMERASES OF BYDV, CARMV, MCMV, RCNMV, AND
CC TCV ARE HIGHLY SIMILAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02986; CAA26726.1; ALT_TERM.
DR PIR; A04208; RRVECV.
DR InterPro; IPR002564; PV_Rdrp.
DR Pfam; PF01615; PV_Rdrp; 1.
KW Transferase; RNA-directed RNA polymerase.
SQ SEQUENCE 867 AA; 97408 MW; 40050A3A24BDABE6 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 867;
Best Local Similarity 44.4%; Pred. No. 63;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
| | | | |
Db 457 WSAWNQFQT 465

RESULT 28
YGCB_ECOLI
ID YGCB_ECOLI STANDARD; PRT; 888 AA.
AC P38036; Q46902;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ygcB.
GN YGCB OR B2761.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
```

RP SEQUENCE OF 1-335 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=91172132; PubMed=2005873;
RA Krone F.A., Westphal G., Schwenn J.D.;
RT "Characterisation of the gene *cysH* and of its product phospho-
adenylsulphate reductase from *Escherichia coli*.";
RL Mol. Gen. Genet. 225:314-319(1991).
RN [3]
RP IDENTIFICATION.
RX MEDLINE=95075659; PubMed=7984428;
RA Borodovsky M., Rudd K.E., Koonin E.V.;
RT "Intrinsic and extrinsic approaches for detecting genes in a
bacterial genome.";
RL Nucleic Acids Res. 22:4756-4767(1994).
CC -!- SIMILARITY: BELONGS TO THE PULE/OUTE/EXEE/XPSE/XCPR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U29579; AAA69271.1; -.
CC EMBL; AE000359; AAC75803.1; -.
CC EMBL; Y07525; -; NOT_ANNOTATED_CDS.
CC EcoGene; EGI2634; YGCB.
CC InterPro; IPR001482; GSPIL_E.
CC InterPro; IPR001650; Helicase_C.
CC Pfam; PF00271; helicase_C; 1.
CC SMART; SM00490; HELICC; 1.
CC PROSITE; PS00662; T2SP_E; FALSE_NEG.
KW Hypothetical protein; Transport; ATP-binding; Complete proteome.
FT NP_BIND 314 321 ATP (POTENTIAL).
FT CONFLICT 118 118 G -> R (IN REF. 2).
FT CONFLICT 334 335 QQ -> PL (IN REF. 2).
SQ SEQUENCE 888 AA; 100544 MW; 4CA3F5371B1BF0F2 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 888;
Best Local Similarity 45.5%; Pred. No. 65;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNE 52
| | | | |
Db 234 WTTNTTFLFNE 244

RESULT 29
Y188_HUMAN
ID Y188_HUMAN STANDARD; PRT; 899 AA.
AC Q14693;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein KIAA0188 (Fragment).
GN KIAA0188.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96281124; PubMed=8724849;
RA Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. V.
RT The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 3:17-24(1996).
CC -!- SIMILARITY: TO HUMAN KIAA0249 AND YEAST SMP2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D80010; BAA11505.1; -.
CC Genew; HGNC:13345; LPIN1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 899 AA; 99756 MW; AF81E5B095035924 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 899;
Best Local Similarity 28.6%; Pred. No. 66;
Matches 4; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXXQTXLXNE 52
| | | | |
Db 563 WFSWGRGNTTIKEE 576

RESULT 30
TOP2_ASFM2
ID TOP2_ASFM2 STANDARD; PRT; 1191 AA.
AC P34203;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA topoisomerase II (EC 5.99.1.3).
GN TOP OR P1192R OR I7R.
OS African swine fever virus (isolate Malawi Lil 20/1) (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10500;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93108435; PubMed=1335084;
RA Baylis S.A., Dixon L.K., Vydelingum S., Smith G.L.;
RT "African swine fever virus encodes a gene with extensive homology to
RT type II DNA topoisomerases.";
RL J. Mol. Biol. 228:1003-1010(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94292916; PubMed=8021596;
RA Dixon L.K., Twigg S.R.F., Baylis S.A., Vydelingum S., Bristow C.,
RA Hammond J.M., Smith G.L.;
RT "Nucleotide sequence of a 55 kbp region from the right end of the
RT genome of a pathogenic African swine fever virus isolate (Malawi
RT Lil20/1).";
RL J. Gen. Virol. 75:1655-1684(1994).
CC -!- FUNCTION: CAN INTRODUCE NEGATIVE SUPERHELICAL TURNS INTO
CC DOUBLE-STRANDED CIRCULAR DNA.
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -!- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z14245; CAA78614.1; -.
CC EMBL; X71982; CAA50820.1; -.
DR PIR; S27329; S27329.
DR HSP; P06786; 1BGW.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR001241; DNA_topoisomII.
DR InterPro; IPR002205; DNA_topoisomIV.
DR Pfam; PF00204; DNA_gyraseB; 1.
DR Pfam; PF00521; DNA_topoisomIV; 1.

DR Pfam; PF02518; HATPase_C; 1.
DR PRINTS; PR00418; TPI2FAMILY.
DR ProDom; PD000742; DNA_topoisoIV; 1.
DR SMART; SM00433; TOP2C; 1.
DR SMART; SM00434; TOP4C; 1.
DR PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW Topoisomerase; Isomerase; DNA-binding; ATP-binding.
FT NP_BIND 142 147 ATP (POTENTIAL).
FT ACT_SITE 799 799 DNA CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 1191 AA; 135070 MW; 62858BFF3F67CE22 CRC64;

Query Match 20.9%; Score 29; DB 1; Length 1191;
Best Local Similarity 50.0%; Pred. No. 84;
Matches 5; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 42 WXXXQTXLXN 51
| | | | |
Db 598 WAKKQTSLAN 607

RESULT 31
IFE2_WHEAT STANDARD; PRT; 209 AA.
AC Q03389;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Eukaryotic translation initiation factor 4E (EIF-4E) (mRNA
DE cap-binding protein) (EIF-(ISO)4F 25 kDa subunit) (EIF-(ISO)4F P28
DE subunit).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054655; PubMed=1385417;
RA Allen M.L., Metz A.M., Timmer R.T., Rhoads R.E., Browning K.S.;
RT "Isolation and sequence of the cDNAs encoding the subunits of the
RT isozyme form of wheat protein synthesis initiation factor 4F.";
RL J. Biol. Chem. 267:23232-23236(1992).
CC -!- FUNCTION: RECOGNIZES AND BINDS THE 7-METHYLGUANOSINE-CONTAINING
CC MRNA "CAP" DURING AN EARLY STEP IN THE INITIATION OF PROTEIN
CC SYNTHESIS AND FACILITATES RIBOSOME BINDING BY INDUCING THE
CC UNWINDING OF THE MRNAS SECONDARY STRUCTURES.
CC -!- SUBUNIT: EIF4F IS A TRIMER COMPOSED OF EIF4E, EIF4G AND EIF4A
CC (WHICH CAN CYCLE IN AND OUT OF THE COMPLEX). IN HIGHER PLANTS TWO
CC ISOFORMS OF EIF4F HAVE BEEN IDENTIFIED, NAMED EIF4F AND
CC EIF(ISO)4F. EIF4F HAS SUBUNITS P220 AND P28, WHEREAS EIF-(ISO)4F
CC HAS SUBUNITS P82 AND P26.
CC -!- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M95818; AAA34295.1; -.
CC HSSP; P07260; 1AP8.
CC InterPro; IPR001040; TIF_eIF_4E.
CC Pfam; PF01652; IF4E; 1.
CC ProDom; PD003697; TIF_eIF_4E; 1.
CC PROSITE; PS00813; IF4E; 1.
KW Initiation factor; Protein biosynthesis; RNA-binding;
KW Multigene family.
SQ SEQUENCE 209 AA; 23522 MW; E6EC9E8557AB2BED CRC64;

QY 42 WXXXQTXLXN 51
| | | | |
Db 598 WAKKQTSLAN 607

Query Match 20.1%; Score 28; DB 1; Length 209;
Best Local Similarity 44.4%; Pred. No. 29;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
| | | | |
Db 39 WTFWYDIQT 47

RESULT 32
IFE2_MAIZE STANDARD; PRT; 216 AA.
AC O81482;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Eukaryotic translation initiation factor 4E (EIF-4E) (EIF4E) (mRNA
DE cap-binding protein) (EIF-(ISO)4F 25 kDa subunit) (EIF-(ISO)4F P28
DE subunit).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root;
RA Manjunath S., Bailey-Serres J.;
RT "Oxygen deprivation stimulates Ca2+ mediated phosphorylation of mRNA
RT cap-binding protein eIF4E in maize roots.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: RECOGNIZES AND BINDS THE 7-METHYLGUANOSINE-CONTAINING
CC MRNA "CAP" DURING AN EARLY STEP IN THE INITIATION OF PROTEIN
CC SYNTHESIS AND FACILITATES RIBOSOME BINDING BY INDUCING THE
CC UNWINDING OF THE MRNAS SECONDARY STRUCTURES.
CC -!- SUBUNIT: EIF4F IS A TRIMER COMPOSED OF EIF4E, EIF4G AND EIF4A
CC (WHICH CAN CYCLE IN AND OUT OF THE COMPLEX). IN HIGHER PLANTS TWO
CC ISOFORMS OF EIF4F HAVE BEEN IDENTIFIED, NAMED EIF4F AND
CC EIF(ISO)4F. EIF4F HAS SUBUNITS P220 AND P28, WHEREAS EIF-(ISO)4F
CC HAS SUBUNITS P82 AND P26.
CC -!- SIMILARITY: BELONGS TO THE EUKARYOTIC INITIATION FACTOR 4E FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF076955; AAC27715.1; -.
CC HSSP; P07260; 1AP8.
CC InterPro; IPR001040; TIF_eIF_4E.
CC Pfam; PF01652; IF4E; 1.
CC ProDom; PD003697; TIF_eIF_4E; 1.
CC PROSITE; PS00813; IF4E; 1.
KW Initiation factor; Protein biosynthesis; RNA-binding;
KW Multigene family.
SQ SEQUENCE 216 AA; 24104 MW; 20FBAEADB206FAC5 CRC64;

Query Match 20.1%; Score 28; DB 1; Length 216;
Best Local Similarity 44.4%; Pred. No. 30;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
| | | | |
Db 46 WTFWYDIQT 54

RESULT 33
PSB5_YEAST STANDARD; PRT; 287 AA.
ID PSB5_YEAST

AC P30656;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteasome component PRE2 precursor (EC 3.4.25.1) (Macropain subunit
DE PRE2) (Proteinase YSCE subunit PRE2) (Multicatalytic endopeptidase
DE complex subunit PRE2).
GN PRE2 OR PRG1 OR DOA3 OR YPR103W OR P8283.10.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93083985; PubMed=1452031;
RA Friedman H., Goebel M., Snyder M.;
RT "A homolog of the proteasome-related RING10 gene is essential for
RT yeast cell growth.";
RL Gene 122:203-206(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RX MEDLINE=93186825; PubMed=8383129;
RA Heinemeyer W., Gruhler A., Moehrl V., Mahe Y., Wolf D.H.;
RT "PRE2, highly homologous to the human major histocompatibility
RT complex-linked RING10 gene, codes for a yeast proteasome subunit
RT necessary for chymotryptic activity and degradation of ubiquitinated
RT proteins.";
RL J. Biol. Chem. 268:5115-5120(1993).
RN [3]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 76-90.
RX MEDLINE=95300795; PubMed=7781614;
RA Chen P., Hochstrasser M.;
RT "Biogenesis, structure and function of the yeast 20S proteasome.";
RL EMBO J. 14:2620-2630(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoerge W.,
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL Nature 387:103-105(1997).
RN [5]
RP SEQUENCE OF 176-287 FROM N.A.
RC STRAIN=FL100;
RX MEDLINE=94217688; PubMed=8164651;
RA Hermann-Ledenmat S., Werner M., Sentenac A., Thuriaux P.;
RT "Suppression of yeast RNA polymerase III mutations by FHL1, a gene
RT coding for a fork head protein involved in rRNA processing.";
RL Mol. Cell. Biol. 14:2905-2913(1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 76-287.
RX MEDLINE=97242404; PubMed=9087403;
RA Groll M., Ditzel L., Lowe J., Stock D., Bochtler M., Bartunik H.D.,
RA Huber R.;
RT "Structure of 20S proteasome from yeast at 2.4-A resolution.";
RL Nature 386:463-471(1997).
CC -!- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX

CC WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH
CC ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT
CC NEUTRAL OR SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT
CC PROTEOLYTIC ACTIVITY. THIS SUBUNIT HAS A CHYMOTRYPSIN-LIKE
CC ACTIVITY.
CC -!- FUNCTION: THIS SUBUNIT IS NECESSARY FOR CHYMOTRYPTIC ACTIVITY AND
CC DEGRADATION OF UBIQUITINATED PROTEINS.
CC -!- CATALYTIC ACTIVITY: Cleavage at peptide bonds with very broad
CC specificity.
CC -!- SUBUNIT: YEAST PROTEASOME SEEMS TO BE COMPOSED OF 14 DIFFERENT
CC SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M96667; AAA34906.1; -;
CC EMBL; X68662; CAA48628.1; -;
CC EMBL; U32445; AAB68073.1; -;
CC EMBL; Z28348; CAA82203.1; ALT_INIT.
CC PIR; JN0446; JN0446.
CC PIR; S29702; S29702.
CC PIR; A45411; A45411.
CC PDB; 1RYP; 15-APR-98.
CC MEROPS; T01.012; -;
CC SGD; S0006307; PRE2.
CC InterPro; IPR000243; Proteasome_B.
CC InterPro; IPR001353; Proteasome_protease.
CC Pfam; PF00227; proteasome; 1.
CC PRINTS; PR00141; PROTEASOME.
CC PROSITE; PS00854; PROTEASOME_B; 1.
KW Proteasome; Hydrolase; Protease; Zymogen; 3D-structure.
FT PROPEP 1 75
FT CHAIN 76 287 PROTEASOME COMPONENT PRE2.
FT VARIANT 126 126 D -> N (IN DOA3-1; DECREASE IN ACTIVITY).
FT CONFLICT 282 287 FNNVIG -> STTLAK (IN REF. 1).
SQ SEQUENCE 287 AA; 31636 MW; DOEBAC611F7A4F37 CRC64;

Query Match 20.1%; Score 28; DB 1; Length 287;
Best Local Similarity 22.2%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 23 PXXGXGXXXXXXXWXXXWXXXQTXL 49
Db 114 PFLGLTMAGGADCQFWETWLGSCRL 140

RESULT 34
NU4M_LATCH
ID NU4M_LATCH STANDARD; PRT; 461 AA.
AC O03173;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN MTND4 OR ND4 OR NADH4.
OS Latimeria chalumnae (Latimeria) (Coelacanth).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Coelacanthiformes; Coelacanthidae; Latimeria.
OX NCBI_TaxID=7897;
RN [1]
RP SEQUENCE FROM N.A.
RA Zardoya R., Meyer A.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U82228; AAC60327.1; -.
DR InterPro; IPR003918; NADHub_oxred4.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR000260; Oxidored_q5_N.
DR Pfam; PF00361; oxidored_q1; 1.
DR Pfam; PF01059; oxidored_q5_N; 1.
DR PRINTS; PR01437; NUOXDRDTASE4.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 461 AA; 52047 MW; 2DEB17101B0AA082 CRC64;

Query Match 20.1%; Score 28; DB 1; Length 461;
Best Local Similarity 36.4%; Pred. No. 60;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | |
Db 389 WSNWTLTMTGL 399

RESULT 35
ETS2_MOUSE
ID ETS2_MOUSE STANDARD; PRT; 468 AA.
AC P15037;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-ETS-2 protein.
GN ETS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89042086; PubMed=2847145;
RA Watson D.K., McWilliams M.J., Lapis P., Lautenberger J.A.,
RA Schweinfest C.W., Papas T.S.;
RT "Mammalian ets-1 and ets-2 genes encode highly conserved proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:7862-7866(1988).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -!- SIMILARITY: CONTAINS 1 POINTED (PNT) DOMAIN.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; J04103; AAA37581.1; -.
DR PIR; C32066; TVMSE2.
DR HSSP; P14921; 2STT.
DR TRANSFAC; T01397; -.
DR MGD; MGI:95456; Ets2.
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ETs.
DR Pfam; PF00178; Ets; 1.
DR Pfam; PF02198; SAM_PNT; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR SMART; SM00251; SAM_PNT; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.

DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS50061; ETS_DOMAIN_3; 1.
KW Proto-oncogene; DNA-binding; Nuclear protein.
FT DOMAIN 87 170 POINTED.
FT DNA_BIND 362 442 ETS-DOMAIN.
SQ SEQUENCE 468 AA; 52827 MW; 5260F3085B7EB831 CRC64;

Query Match 20.1%; Score 28; DB 1; Length 468;
Best Local Similarity 30.8%; Pred. No. 61;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | |
Db 114 WLLWATNEFSLVN 126

RESULT 36
ETS2_HUMAN
ID ETS2_HUMAN STANDARD; PRT; 469 AA.
AC P15036;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-ets-2 protein.
GN ETS2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89042086; PubMed=2847145;
RA Watson D.K., McWilliams M.J., Lapis P., Lautenberger J.A.,
RA Schweinfest C.W., Papas T.S.;
RT "Mammalian ets-1 and ets-2 genes encode highly conserved proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:7862-7866(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Zimmermann W.W.K., Korenberg J., Rosenthal A., Schattevoy R.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schilhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
RN [4]
RP SEQUENCE OF 399-469 FROM N.A.
RX MEDLINE=91067187; PubMed=2250910;
RA Watson D.K., Mavrothalassitis G.J., Jorczyk C.L., Smyth F.E.,
RA Papas T.S.;
RT "Molecular organization and differential polyadenylation sites of the
RT human ETS2 gene.";
RL Oncogene 5:1521-1527(1990).
RN [5]
RP SEQUENCE OF 324-469 FROM N.A.
RX MEDLINE=86042652; PubMed=2997781;
RA Watson D.K., McWilliams-Smith M.J., Nunn M.F., Duesberg P.H.,
RA O'Brien S.J., Papas T.S.;
RT "The ets sequence from the transforming gene of avian
RT erythroblastosis virus, E26, has unique domains on human chromosomes
RT 11 and 21: both loci are transcriptionally active.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 82:7294-7298(1985).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -!- SIMILARITY: CONTAINS 1 POINTED (PNT) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04102; AAA52412.1; -.
DR EMBL; AF017257; AAB94057.1; -.
DR EMBL; AP001732; BAA95514.1; -.
DR EMBL; AL163278; CAB90468.1; -.
DR EMBL; X55181; CAA38966.1; -.
DR EMBL; M11922; AAA52411.1; -.
DR PIR; B32066; TVHUE2.
DR HSSP; P14921; 2STT.
DR TRANSFAC; T00113; -.
DR Genew; HGNC:3489; ETS2.
DR MIM; 164740; -.
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ETS.
DR InterPro; IPR003118; SAM_PNT.
DR Pfam; PF00178; Ets; 1.
DR Pfam; PF02198; SAM_PNT; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR SMART; SM00251; SAM_PNT; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS0061; ETS_DOMAIN_3; 1.
KW Proto-oncogene; DNA-binding; Nuclear protein.
FT DOMAIN 87 170 POINTED.
FT DNA_BIND 363 443 ETS-DOMAIN.
FT SEQUENCE 469 AA; 53001 MW; 5944EC4B5AAB553E CRC64;

Query Match 20.1%; Score 28; DB 1; Length 469;
Best Local Similarity 30.8%; Pred. No. 61;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
Db 114 WLLWATNEFSLVN 126

RESULT 37
ET2B_XENLA STANDARD; PRT; 472 AA.
AC Q91712;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-ETS-2B protein.
GN ETS2B OR ETS-2B.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=90356411; PubMed=2201951;
RA Wolff C.M., Stiegler P., Baltzinger M., Meyer D., Ghysdael J.,
RA Stehelin D., Bfort N., Remy P.;
RT "Isolation of two different c-ets-2 proto-oncogenes in Xenopus
RT laevis.";
RL Nucleic Acids Res. 18:4603-4604(1990).
RN [2]
```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=92088972; PubMed=1751411;
RA Wolff C.M., Stiegler P., Baltzinger M., Meyer D., Ghysdael J.,
RA Stehelin D., Bfort N., Remy P.;
RT "Cloning, sequencing, and expression of two Xenopus laevis c-ets-2
RT protooncogenes.";
RL Cell Growth Differ. 2:447-456(1991).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -!- SIMILARITY: CONTAINS 1 POINTED (PNT) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X52635; CAA36860.1; -.
DR HSSP; P14921; 2STT.
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ETS.
DR InterPro; IPR003118; SAM_PNT.
DR Pfam; PF00178; Ets; 1.
DR Pfam; PF02198; SAM_PNT; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR SMART; SM00251; SAM_PNT; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS0061; ETS_DOMAIN_3; 1.
KW DNA-binding; Nuclear protein.
FT DOMAIN 87 170 POINTED.
FT DNA_BIND 366 446 ETS-DOMAIN.
FT SEQUENCE 472 AA; 53928 MW; E98AFC77F4D446DD CRC64;

Query Match 20.1%; Score 28; DB 1; Length 472;
Best Local Similarity 30.8%; Pred. No. 61;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
Db 114 WLLWAAKEFSLEN 126

RESULT 38
AMY2_SALTY STANDARD; PRT; 494 AA.
AC P26613;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytoplasmic alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
GN AMYA OR STM1963.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SJW1103;
RX MEDLINE=93015717; PubMed=1400215;
RA Raha M., Kawagishi I., Mueller V., Kihara M., Macnab R.M.;
RT "Escherichia coli produces a cytoplasmic alpha-amylase, AmyA.";
RL J. Bacteriol. 174:6644-6652(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
```

RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE OF 1-6 FROM N.A.
RC STRAIN=SJW1103;
RX MEDLINE=92407478; PubMed=1527488;
RA Kawagishi I., Mueller V., Williams A.W., Irikura V.M., Macnab R.M.;
RT "Subdivision of flagellar region III of the *Escherichia coli* and
RT *Salmonella typhimurium* chromosomes and identification of two
RT additional flagellar genes.";
RL J. Gen. Microbiol. 138:1051-1065(1992).
RN [4]
RP SEQUENCE OF 476-494 FROM N.A.
RX MEDLINE=93381452; PubMed=8371104;
RA Raha M., Kihara M., Kawagishi I., Macnab R.M.;
RT "Organization of the *Escherichia coli* and *Salmonella typhimurium*
RT chromosomes between flagellar regions IIIa and IIib, including a
RT large non-coding region.";
RL J. Gen. Microbiol. 139:1401-1407(1993).
CC -|- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L01643; AAA27110.1; -;
DR EMBL; AE008787; AAL20875.1; -;
DR EMBL; M85241; AAA27079.1; -;
DR EMBL; L13280; AAA71970.1; -;
DR PIR; B45738; B45738.
DR HSSP; P06278; 1VJS.
DR StyGene; SG10011; amya.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Complete proteome.
FT ACT_SITE 235 235 BY SIMILARITY.
FT ACT_SITE 239 239 BY SIMILARITY.
FT ACT_SITE 332 332 BY SIMILARITY.
FT CONFLICT 462 462 L -> S (IN REF. 1).
SQ SEQUENCE 494 AA; 56522 MW; 5C1E862FEDD5E47C CRC64;

Query Match 20.1%; Score 28; DB 1; Length 494;
Best Local Similarity 44.4%; Pred. No. 64;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
Db 219 WARWVMEQT 227

RESULT 39
AMY2_ECOLI
ID AMY2_ECOLI STANDARD; PRT; 495 AA.
AC P26612; P78072;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytoplasmic alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
GN AMYA OR B1927.

OS *Escherichia coli*.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC *Escherichia*.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JA11;
RX MEDLINE=93015717; PubMed=1400215;
RA Raha M., Kawagishi I., Mueller V., Kihara M., Macnab R.M.;
RT "Escherichia coli produces a cytoplasmic alpha-amylase, AmyA.";
RL J. Bacteriol. 174:6644-6652(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Sivasubram S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the *Escherichia coli* K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
RN [4]
RP SEQUENCE OF 1-5 FROM N.A.
RC STRAIN=JA11;
RX MEDLINE=92407478; PubMed=1527488;
RA Kawagishi I., Mueller V., Williams A.W., Irikura V.M., Macnab R.M.;
RT "Subdivision of flagellar region III of the *Escherichia coli* and
RT *Salmonella typhimurium* chromosomes and identification of two
RT additional flagellar genes.";
RL J. Gen. Microbiol. 138:1051-1065(1992).
RN [5]
RP SEQUENCE OF 475-495 FROM N.A.
RC STRAIN=JA11;
RX MEDLINE=93381452; PubMed=8371104;
RA Raha M., Kihara M., Kawagishi I., Macnab R.M.;
RT "Organization of the *Escherichia coli* and *Salmonella typhimurium*
RT chromosomes between flagellar regions IIIa and IIib, including a
RT large non-coding region.";
RL J. Gen. Microbiol. 139:1401-1407(1993).
CC -|- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L01642; AAA23810.1; -;
DR EMBL; AE000285; AAC74994.1; -;
DR EMBL; D90833; BAA15755.1; -;
DR EMBL; M85240; -; NOT_ANNOTATED_CDS.
DR EMBL; L13279; AAA82575.1; -;
DR PIR; A45738; A45738.
DR HSSP; P06278; 1VJS.

```
DR EcoGene; EG11387; amYA.
DR InterPro; IPR000461; Alpha_amylase.
DR Pfam; PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Complete proteome.
FT ACT_SITE 235 235 BY SIMILARITY.
FT ACT_SITE 239 239 BY SIMILARITY.
FT ACT_SITE 332 332 BY SIMILARITY.
FT CONFLICT 19 20 KL -> SS (IN REF. 1).
FT CONFLICT 109 109 A -> V (IN REF. 1).
FT CONFLICT 149 149 Q -> E (IN REF. 1).
FT CONFLICT 234 234 L -> I (IN REF. 1).
SQ SEQUENCE 495 AA; 56639 MW; 26AFF6797DDA54D6 CRC64;

Query Match 20.1%; Score 28; DB 1; Length 495;
Best Local Similarity 44.4%; Pred. No. 64;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
   | | | |
Db 219 WARWMEQT 227

RESULT 40
REPR_STRAG STANDARD; PRT; 496 AA.
AC P18629;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Repr protein.
GN REPR.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90370459; PubMed=2118624;
RA Brantl S., Behnke D., Alonso J.C.;
RT "Molecular analysis of the replication region of the conjugative
RT Streptococcus agalactiae plasmid pIP501 in Bacillus subtilis.
RT Comparison with plasmids pAM beta 1 and pSM19035."
RL Nucleic Acids Res. 18:4783-4790(1990).
CC -!- FUNCTION: ESSENTIAL FOR REPLICATION.
CC -!- SIMILARITY: TO THE REP PROTEINS FROM OTHER PLASMIDS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X17655; CAA35647.1; -
DR PIR; S11211; S11211.
KW Plasmid; DNA replication; DNA-binding.
FT DNA_BIND 120 141 POTENTIAL.
SQ SEQUENCE 496 AA; 57408 MW; 43DFB9B721CA2098 CRC64;

Query Match 20.1%; Score 28; DB 1; Length 496;
Best Local Similarity 44.4%; Pred. No. 64;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
   | | | |
Db 216 WQDWSFKQT 224

RESULT 41
REPS_STRPY STANDARD; PRT; 496 AA.
ID REPS_STRPY
```

```
AC P14752; Q54852;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE RePS protein.
GN REPS.
OS Streptococcus pyogenes.
OG Plasmid pSM19035, Plasmid pMD101, and Plasmid pBT233.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX PLASMID=pSM19035;
RX MEDLINE=90098785; PubMed=26900001;
RA Brantl S., Nowak A., Behnke D., Alonso J.C.;
RT "Revision of the nucleotide sequence of the Streptococcus pyogenes
RT plasmid pSM19035 reps gene."
RL Nucleic Acids Res. 17:10110-10110(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX PLASMID=pSM19035;
RA Sorokin A.V., Khazak V.E.;
RL (In) Butler L.O., Harwood C., Moseley B.E.B. (eds.);
RL Genetic transformation and expression, pp.269-281, Intercept,
RL Andover (1989).
RN [3]
RP SEQUENCE FROM N.A.
RX PLASMID=pBT233;
RX MEDLINE=94123985; PubMed=8293991;
RA Ceglowski P., Boitsov A., Chai S., Alonso J.C.;
RT "Analysis of the stabilization system of pSM19035-derived plasmid
RT pBT233 in Bacillus subtilis."
RL Gene 136:1-12(1993).
RN [4]
RP SEQUENCE FROM N.A.
RX PLASMID=pMD101;
RX MEDLINE=94320784; PubMed=8093174;
RA Ceglowski P., Alonso J.C.;
RT "Gene organization of the Streptococcus pyogenes plasmid pDB101:
RT sequence analysis of the orf eta-cops region."
RL Gene 145:33-39(1994).
RN [5]
RP SEQUENCE OF 436-496 FROM N.A.
RX PLASMID=pSM19035;
RX MEDLINE=91066902; PubMed=2123523;
RA Sorokin A.V., Khazak V.E.;
RT "Expression unit in the region of replication initiation in the
RT streptococcal plasmid pSM19035."
RL Mol. Biol. (Mosk) 24:993-1000(1990).
CC -!- FUNCTION: ESSENTIAL FOR REPLICATION.
CC -!- SIMILARITY: TO THE REP PROTEINS FROM OTHER PLASMIDS.
CC -!- CAUTION: REF.1 SEQUENCE REVISES THAT PUBLISHED IN REF.2.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X16803; CAA34721.1; -
DR EMBL; X64695; CAA45928.1; -
DR EMBL; X66468; CAA47078.1; -
DR EMBL; X61167; CAA43478.1; -
DR PIR; S06925; S06925.
DR PIR; S45072; S45072.
KW Plasmid; DNA replication; DNA-binding.
FT DNA_BIND 120 141 POTENTIAL.
FT CONFLICT 438 438 G -> V (IN REF. 5).
SQ SEQUENCE 496 AA; 57373 MW; 269712933C55E010 CRC64;
```



```
Query Match      20.1%; Score 28; DB 1; Length 496;
Best Local Similarity 44.4%; Pred. No. 64;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQT 47
Db 216 WQDWSFKQT 224

RESULT 42
Y110_TREPA
ID Y110_TREPA STANDARD; PRT; 577 AA.
AC O83148;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0110.
GN TP0110.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT Spirochete."
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001196; AAC65108.1; -.
DR TIGR; TP0110; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 577 AA; 65149 MW; 55735A7173C3FF98 CRC64;

Query Match      20.1%; Score 28; DB 1; Length 577;
Best Local Similarity 36.4%; Pred. No. 73;
Matches 4; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
Db 357 WALWSTACKRL 367

RESULT 43
PLBL_TORDE
ID PLBL_TORDE STANDARD; PRT; 649 AA.
AC Q11121;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Lysophospholipase precursor (EC 3.1.1.5) (Phospholipase B).
OS Torulaspora delbrueckii (Yeast) (Saccharomycetes rosei).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Torulaspora.
OX NCBI_TaxID=4950;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=YL-32;
```

```
RX MEDLINE=95095025; PubMed=8001766;
RA Watanabe Y., Yashiki Y., Sultana G.N., Maruyama M., Kangawa K.,
RA Tamai Y.;
RT "Cloning and sequencing of phospholipase B gene from the yeast
RT Torulaspora delbrueckii."
RL FEMS Microbiol. Lett. 124:29-34(1994).
CC -!- FUNCTION: CATALYZES THE RELEASE OF FATTY ACIDS FROM
CC LYSOPHOSPHOLIPIDS.
CC -!- CATALYTIC ACTIVITY: 2-lysophosphatidylcholine + H(2)O =
CC glycerophosphocholine + a fatty acid anion.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: TO OTHER FUNGAL LYSOPHOSPHOLIPASES AND TO YEAST
CC SPO1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D32134; BAA06860.1; -.
DR InterPro; IPR002642; PLAC.
DR Pfam; PF01735; PLA2_B; 1.
DR SMART; SM00022; PLAC; 1.
KW Lipid degradation; Hydrolase; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 649
FT DOMAIN 612 617
FT CARBOHYD 32 32
FT CARBOHYD 51 51
FT CARBOHYD 77 77
FT CARBOHYD 90 90
FT CARBOHYD 121 121
FT CARBOHYD 158 158
FT CARBOHYD 168 168
FT CARBOHYD 213 213
FT CARBOHYD 275 275
FT CARBOHYD 343 343
FT CARBOHYD 386 386
FT CARBOHYD 457 457
FT CARBOHYD 487 487
FT CARBOHYD 511 511
FT CARBOHYD 539 539
FT CARBOHYD 563 563
FT CARBOHYD 580 580
SQ SEQUENCE 649 AA; 71110 MW; CC6901D168D01459 CRC64;

Query Match      20.1%; Score 28; DB 1; Length 649;
Best Local Similarity 30.8%; Pred. No. 81;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
Db 156 WNNWTSVQDIVNN 168

RESULT 44
CTPB_MYCLE
ID CTPB_MYCLE STANDARD; PRT; 750 AA.
AC P46840;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cation-transporting p-type ATPase B (EC 3.6.3.-).
GN CTPB OR ML2000.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
```


RX MEDLINE=96059637; PubMed=7476188;
RA Fsihi H., Cole S.T.;
RT "The Mycobacterium leprae genome: systematic sequence analysis
RT identifies key catabolic enzymes, ATP-dependent transport systems and
RT a novel polA locus associated with genomic variability.";
RL Mol. Microbiol. 16:909-919(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC (E1-E2 ATPASES). SUBFAMILY IB.
CC -!- SIMILARITY: CONTAINS 1 HMA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46257; CAA86363.1; -.
DR EMBL; AL583924; CAC30955.1; -.
DR HSSP; P04129; IAFJ.
DR Leproma; ML2000; -.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR001756; Cu_ATPase.
DR InterPro; IPR001934; HeavyMe_transpt.
DR InterPro; IPR001454; Hlgnaase/hydrase.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00403; HMA; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00943; CUATPASE.
DR PROSITE; PS00154; ATPASE_E1_E2; 1.
DR PROSITE; PS01047; HMA_1; 1.
DR PROSITE; PS50846; HMA_2; 1.
DR Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
KW Metal-binding; Complete proteome.
FT TRANSMEM 104 124 POTENTIAL.
FT TRANSMEM 129 149 POTENTIAL.
FT TRANSMEM 167 187 POTENTIAL.
FT TRANSMEM 200 220 POTENTIAL.
FT TRANSMEM 360 380 POTENTIAL.
FT TRANSMEM 389 409 POTENTIAL.
FT TRANSMEM 471 491 POTENTIAL.
FT TRANSMEM 500 520 POTENTIAL.
FT TRANSMEM 547 567 POTENTIAL.
FT TRANSMEM 663 683 POTENTIAL.
FT TRANSMEM 693 713 POTENTIAL.
FT TRANSMEM 715 735 POTENTIAL.
FT DOMAIN 18 81 HMA.
FT MOD_RES 445 445 PHOSPHORYLATION (PROBABLE).
FT METAL 28 28 POTENTIAL.
FT METAL 31 31 POTENTIAL.
FT CONFLICT 544 548 EGSER -> RRRIT (IN REF. 1).
SQ SEQUENCE 750 AA; 78100 MW; 1D58477D4A69B00D CRC64;

Query Match 20.1%; Score 28; DB 1; Length 750;

Best Local Similarity 30.8%; Pred. No. 93;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXLXN 51
| | : | |
Db 146 WAAWPFHRVALRN 158
RESULT 45
RB_RAT RB_RAT STANDARD; PRT; 899 AA.
AC P33568; Q63527;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Retinoblastoma-associated protein (PPI05) (RB) (Fragment).
GN RB1 OR RB-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95394362; PubMed=7665085;
RA Esumi M., Idutsu T., Kinugasa S., Ohno M., Nakabayashi H., Ikeda T.,
RA Shikata T.;
RT "Isolation and sequence polymorphism of a rat retinoblastoma (RB)
RT cDNA";
RL Gene 161:231-235(1995).
RN [2]
RP SEQUENCE OF 785-899 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=93181172; PubMed=8441612;
RA Roy N.K., Ballesteros A., Garte S.J.;
RT "Cloning and sequence of the rat retinoblastoma (Rb) gene cDNA.";
RL Nucleic Acids Res. 21:170-170(1993).
RN [3]
RP INTERACTION WITH DNMT1.
RX MEDLINE=20347723; PubMed=10888886;
RA Robertson K.D., Ait-Si-Ali S., Yokochi T., Wade P.A., Jones P.L.,
RA Wolffe A.P.;
RT "DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses
RT transcription from E2F-responsive promoters.";
RL Nat. Genet. 25:338-342(2000).
CC -!- FUNCTION: PROBABLY ACTS AS A REGULATOR OF OTHER GENES. FORMS A
CC COMPLEX WITH ADENOVIRUS E1A AND WITH SV40 LARGE T ANTIGEN. ACTS AS
CC A TUMOR SUPPRESSOR. MAY BIND AND MODULATE FUNCTIONALLY CERTAIN
CC CELLULAR PROTEINS WITH WHICH T AND E1A COMPETE FOR POCKET BINDING.
CC POTENT INHIBITOR OF E2F-MEDIATED TRANS-ACTIVATION.
CC -!- SUBUNIT: INTERACTS WITH TRANSCRIPTION FACTOR E2F1 AND WITH DNMT1
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: PHOSPHORYLATED FROM S TO M PHASE OF THE CELL CYCLE AND IS
CC DEPHOSPHORYLATED IN G1. T, BUT NOT E1A, BINDS ONLY TO THE
CC UNPHOSPHORYLATED FORM (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE RETINOBLASTOMA PROTEIN (RB) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D25233; BAA04958.1; -.
DR EMBL; L07126; AAA42090.1; -.
DR PIR; S35544; S35544.
DR HSSP; P06400; IGUX.
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR002720; RB_A.
DR InterPro; IPR002719; RB_B.

DR Pfam; PF01857; RB_B; 1.
DR Pfam; PF01858; RB_A; 1.
DR SMART; SM00385; CYCLIN; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; Anti-oncogene.
FT NON_TER 1
FT CONFLICT 845 846 KL -> TW (IN REF. 2).
SQ SEQUENCE 899 AA; 102872 MW; A6D837E29730FA73 CRC64;

Query Match 20.1%; Score 28; DB 1; Length 899;
Best Local Similarity 42.9%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNEXXL 55
Db 652 WTLFQHTLENEYEL 665

RESULT 46
RB_MOUSE
ID RB_MOUSE STANDARD; PRT; 921 AA.
AC P13405;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Retinoblastoma-associated protein (PP105) (RB).
GN RB1 OR RB-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89367271; PubMed=2671991;
RA Bernards R., Schackleford G.M., Gerber M.R., Horowitz J.M.,
RA Friend S.H., Schartl M., Bogenmann E., Rapaport J., McGee T.,
RA Dryja T., Weinberg R.A.;
RT "Structure and expression of the murine retinoblastoma gene and
RT characterization of its encoded protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6474-6478(1989).
RN [2]
RP INTERACTION WITH DNMT1.
RX MEDLINE=20347723; PubMed=10888886;
RA Robertson K.D., Ait-Si-Ali S., Yokochi T., Wade P.A., Jones P.L.,
RA Wolffe A.P.;
RT "DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses
RT transcription from E2F-responsive promoters.";
RL Nat. Genet. 25:338-342(2000).
CC -!- FUNCTION: PROBABLY ACTS AS A REGULATOR OF OTHER GENES. FORMS A
CC COMPLEX WITH ADENOVIRUS E1A AND WITH SV40 LARGE T ANTIGEN. ACTS AS
CC A TUMOR SUPPRESSOR. MAY BIND AND MODULATE FUNCTIONALLY CERTAIN
CC CELLULAR PROTEINS WITH WHICH T AND E1A COMPETE FOR POCKET BINDING.
CC POTENT INHIBITOR OF E2F-MEDIATED TRANS-ACTIVATION.
CC -!- SUBUNIT: INTERACTS WITH TRANSCRIPTION FACTOR E2F1 AND WITH DNMT1
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: PHOSPHORYLATED FROM S TO M PHASE OF THE CELL CYCLE AND IS
CC DEPHOSPHORYLATED IN G1. T, BUT NOT E1A, BINDS ONLY TO THE
CC UNPHOSPHORYLATED FORM.
CC -!- SIMILARITY: BELONGS TO THE RETINOBLASTOMA PROTEIN (RB) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M26391; AAA39964.1; -.
DR PIR; A33718; A33718.
DR HSSP; P06400; 1GUX.
DR MGD; MGI:97874; Rb1.

DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR002720; RB_A.
DR InterPro; IPR002719; RB_B.
DR Pfam; PF01857; RB_B; 1.
DR Pfam; PF01858; RB_A; 1.
DR SMART; SM00385; CYCLIN; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; Anti-oncogene.
FT DOMAIN 9 12 POLY-ALA.
FT DOMAIN 14 22 POLY-PRO.
SQ SEQUENCE 921 AA; 105338 MW; 149ED65A07BC6495 CRC64;

Query Match 20.1%; Score 28; DB 1; Length 921;
Best Local Similarity 42.9%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNEXXL 55
Db 674 WTLFQHTLQNEYEL 687

RESULT 47
RB_HUMAN
ID RB_HUMAN STANDARD; PRT; 928 AA.
AC P06400; P78499;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Retinoblastoma-associated protein (PP110) (P105-RB) (RB).
GN RB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88014238; PubMed=3657987;
RA Lee W.-H., Shew J.-Y., Hong F.D., Sery T.W., Donoso L.A., Young L.-J.,
RA Bookstein R., Lee E.Y.-H.P.;
RT "The retinoblastoma susceptibility gene encodes a nuclear
RT phosphoprotein associated with DNA binding activity.";
RL Nature 329:642-645(1987).
RN [2]
RP REVISIONS.
RX MEDLINE=87149066; PubMed=3823889;
RA Lee W.-H., Bookstein R., Hong F.D., Young L.-J., Shew J.-Y.,
RA Lee E.Y.-H.P.;
RT "Human retinoblastoma susceptibility gene: cloning, identification,
RT and sequence.";
RL Science 235:1394-1399(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88097427; PubMed=3480530;
RA Friend S.H., Horowitz J.M., Gerber M.R., Wang X.-F., Bogenmann E.,
RA Li F.P., Weinberg R.A.;
RT "Deletions of a DNA sequence in retinoblastomas and mesenchymal
RT tumors: organization of the sequence and its encoded protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:9059-9063(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=94063891; PubMed=7902321;
RA Toguchida J., McGee T.L., Ciccarelli J.A., Eagle J.R., Yandell D.W.,
RA Dryja T.P.;
RT "Complete genomic sequence of the human retinoblastoma susceptibility
RT gene.";
RL Genomics 17:535-543(1993).

RA [6] SEQUENCE FROM N.A.
RP TISSUE-Carcinoma;
RX MEDLINE=92319557; PubMed=1352398;
RA Hogg A., Onadim Z., Baird P.N., Cowell J.K.;
RT "Detection of heterozygous mutations in the RB1 gene in
RT retinoblastoma patients using single-strand conformation polymorphism
RT analysis and polymerase chain reaction sequencing.";
RL Oncogene 7:1445-1451(1992).
RN [7]
RP SEQUENCE OF 1-45 FROM N.A.
RX MEDLINE=89239464; PubMed=2717184;
RA T'Ang A., Wu K.J., Hashimoto T., Liu W.Y., Takahashi R., Shi X.H.,
RA Mihara K., Zhang F.H., Chen Y.Y., Du C., Qian J., Lin Y.G.,
RA Murphree A.L., Qiu W.R., Thompson T., Benedict W.F., Fung Y.K.T.;
RT "Genomic organization of the human retinoblastoma gene";
RL Oncogene 4:401-407(1989).
RN [8]
RP COMPLEX WITH SV40 LARGE T ANTIGEN.
RX MEDLINE=88270506; PubMed=2839300;
RA Decaprio J.A., Ludlow J.W., Figge J., Shew J.-Y., Huang C.-M.,
RA Lee W.-H., Marsilio E., Paucha E., Livingston D.M.;
RT "SV40 large tumor antigen forms a specific complex with the product of
RT the retinoblastoma susceptibility gene.";
RL Cell 54:275-283(1988).
RN [9]
RP INTERACTION WITH DNMT1.
RX MEDLINE=20347723; PubMed=10888886;
RA Robertson K.D., Ait-Si-Ali S., Yokochi T., Wade P.A., Jones P.L.,
RA Wolffe A.P.;
RT "DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses
RT transcription from E2F-responsive promoters.";
RL Nat. Genet. 25:338-342(2000).
RN [10]
RP PHOSPHORYLATION SITES.
RX MEDLINE=92097548; PubMed=1756735;
RA Lees J.A., Buchkovich K.J., Marshak D.R., Anderson C.W., Harlow E.;
RT "The retinoblastoma protein is phosphorylated on multiple sites by
RT human cdc2.";
RL EMBO J. 10:4279-4290(1991).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 378-562.
RX MEDLINE=97290453; PubMed=9145110;
RA Kim H.Y., Cho Y.;
RT "Structural similarity between the pocket region of retinoblastoma
RT tumour suppressor and the cyclin-box.";
RL Nat. Struct. Biol. 4:390-395(1997).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 380-785.
RX MEDLINE=98154728; PubMed=9495340;
RA Lee J.O., Russo A.A., Pavletich N.P.;
RT "Structure of the retinoblastoma tumour-suppressor pocket domain bound
RT to a peptide from HPV E7.";
RL Nature 391:859-865(1998).
RN [13]
RP VARIANT RB LEU-567.
RX MEDLINE=90081757; PubMed=2594029;
RA Yandell D.W., Campbell T.A., Dayton S.H., Petersen R., Walton D.,
RA Little J.B., McConkie-Rosell A., Buckley E., Dryja T.;
RT "Oncogenic point mutations in the human retinoblastoma gene: their
RT application to genetic counseling.";
RL New Engl. J. Med. 321:1689-1695(1989).
RN [14]
RP VARIANT RB TRP-661.
RX MEDLINE=92335261; PubMed=1352883;
RA Onadim Z., Hogg A., Baird P.N., Cowell J.K.;
RT "Oncogenic point mutations in exon 20 of the RB1 gene in families
RT showing incomplete penetrance and mild expression of the
RT retinoblastoma phenotype.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:6177-6181(1992).
RN [15]
RP VARIANT RB ARG-457.
RX MEDLINE=93348271; PubMed=8346255;

RA Hogg A., Bia B., Onadim Z., Cowell J.K.;
RT "Molecular mechanisms of oncogenic mutations in tumors from patients
RT with bilateral and unilateral retinoblastoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7351-7355(1993).
RN [16]
RP VARIANT RB ARG-530.
RX MEDLINE=95219415; PubMed=7704558;
RA Cowell J.K., Smith T., Bia B.;
RT "Frequent constitutional C to T mutations in CGA-arginine codons in
RT the RB1 gene produce premature stop codons in patients with bilateral
RT (hereditary) retinoblastoma.";
RL Eur. J. Hum. Genet. 2:281-290(1994).
RN [17]
RP VARIANTS RB ASN-480 DEL AND TRP-661.
RX MEDLINE=95012220; PubMed=7927327;
RA Lohmann D.R., Brandt B., Hoepping W., Passarge E., Horsthenke B.;
RT "Distinct RB1 gene mutations with low penetrance in hereditary
RT retinoblastoma.";
RL Hum. Genet. 94:349-354(1994).
RN [18]
RP VARIANTS RB GLN-72; TYR-549 AND LYS-803.
RX MEDLINE=96187126; PubMed=8605116;
RA Liu Z., Song Y., Bia B., Cowell J.K.;
RT "Germline mutations in the RB1 gene in patients with hereditary
RT retinoblastoma.";
RL Genes Chromosomes Cancer 14:277-284(1995).
RN [19]
RP VARIANTS RB THR-185; PRO-635; GLU-654 AND PRO-685.
RX MEDLINE=95315934; PubMed=7795591;
RA Blanquet V., Turleau C., Gross-Morand M.S., Senamaud-Beaufort C.,
RA Doz F., Besmond C.;
RT "Spectrum of germline mutations in the RB1 gene: a study of 232
RT patients with hereditary and non hereditary retinoblastoma.";
RL Hum. Mol. Genet. 4:383-388(1995).
RN [20]
RP VARIANTS RB GLY-358; PRO-657 AND TRP-661.
RX MEDLINE=96372810; PubMed=8776589;
RA Van Orsouw N.J., Li D., van der Vlies P., Scheffer H., Eng C.,
RA Buys C.H.C.M., Li F.P., Vijg J.;
RT "Mutational scanning of large genes by extensive PCR multiplexing and
RT two-dimensional electrophoresis: application to the RB1 gene.";
RL Hum. Mol. Genet. 5:755-761(1996).
RN [21]
RP VARIANTS RB ASP-137 AND TRP-661.
RX MEDLINE=97456418; PubMed=9311732;
RA Lohmann D.R., Gerick M., Brandt B., Oelschlaeger U., Lorenz B.,
RA Passarge E., Horsthenke B.;
RT "Constitutional RB1-gene mutations in patients with isolated
RT unilateral retinoblastoma.";
RL Am. J. Hum. Genet. 61:282-294(1997).
RN [22]
RP VARIANT RB GLN-447.
RX MEDLINE=97285179; PubMed=9140452;
RA Mateu E., Sanchez F., Najera C., Beneyto M., Castell V., Hernandez M.,
RA Serra I., Prieto F.;
RT "Genetics of retinoblastoma: a study.";
RL Cancer Genet. Cytogenet. 95:40-50(1997).
RN [23]
RP VARIANTS RB LEU-567; ARG-712 AND PRO-662.
RX Yilmaz S., Horsthenke B., Lohmann D.R.;
RA "Twelve novel RB1 gene mutations in patients with hereditary
RT retinoblastoma.";
RL Hum. Mutat. 12:434-434(1998).
RN [24]
RP VARIANT RB GLU-310.
RX MEDLINE=99138683; PubMed=9973307;
RA Klutz M., Horsthenke B., Lohmann D.R.;
RT "RB1 gene mutations in peripheral blood DNA of patients with isolated
RT unilateral retinoblastoma.";
RL Am. J. Hum. Genet. 64:667-668(1999).
RN [25]
RP VARIANTS RB GLY-500 AND GLU-616.
RX MEDLINE=21415586; PubMed=11524739;

RA Yu Y.S., Kim I.-J., Ku J.-L., Park J.-G.;
RT "Identification of four novel RBL germline mutations in Korean
retinoblastoma patients."
RL Hum. Mutat. 18:252-252(2001).

Query Match 20.1%; Score 28; DB 1; Length 928;
Best Local Similarity 42.9%; Pred. No. 1.le+02;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNEXXL 55
| | | | | | | | | |
Db 681 WTLFQHTLQNEYEL 694

RESULT 48

TSPL_HUMAN
ID TSPL_HUMAN STANDARD; PRT; 1170 AA.
AC P07996;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSPL OR TSP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Endothelial cells;
RX MEDLINE=87057617; PubMed=2430973;
RA Lawler J., Hynes R.O.;
RT "The structure of human thrombospondin, an adhesive glycoprotein with
multiple calcium-binding sites and homologies with several different
proteins.";
RL J. Cell Biol. 103:1635-1648(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89139590; PubMed=2918029;
RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
RA Baumgartel D.M., Rotwein P., Frazier W.A.;
RT "Complete thrombospondin mRNA sequence includes potential regulatory
sites in the 3' untranslated region.";
RL J. Cell Biol. 108:729-736(1989).
RN [3]
RP SEQUENCE OF 1-397 FROM N.A.
RX MEDLINE=87157592; PubMed=3030396;
RA Kobayashi S., Eden-McCutchan F., Framson P., Bornstein P.;
RT "Partial amino acid sequence of human thrombospondin as determined by
analysis of cDNA clones: homology to malarial circumsporozoite
proteins.";
RL Biochemistry 25:8418-8425(1986).
RN [4]
RP SEQUENCE OF 1-374 FROM N.A.
RX MEDLINE=86287276; PubMed=3461443;
RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RT "Characterization of a cDNA encoding the heparin and collagen binding
domains of human thrombospondin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
RN [5]
RP SEQUENCE OF 1-166 FROM N.A.
RX MEDLINE=89291870; PubMed=2544587;
RA Laherty C.D., Gierman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin
gene. DNA sequences within the first intron increase transcription.";
RL J. Biol. Chem. 264:11222-11227(1989).
RN [6]
RP SEQUENCE OF 1028-1170 FROM N.A.
RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN [7]
RP CARBOHYDRATE-LINKAGE SITES W-385; S-394; W-438; W-441; T-450; W-498
AND T-507.

RC TISSUE=Platelet;
RX MEDLINE=21125860; PubMed=11067851;
RA Hofsteenge J., Huwiler K.G., Macek B., Hess D., Lawler J.,
RA Mosher D.F., Peter-Katalinic J.;
RT "C-mannosylation and O-fucosylation of the thrombospondin type 1
module.";
RL J. Biol. Chem. 276:6485-6498(2001).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
laminin, type V collagen and integrins alpha-V/beta-1, alpha-
V/beta-3 and alpha-IIB/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 TSP TYPE-1 DOMAINS.
CC -!- SIMILARITY: CONTAINS 7 TSP TYPE-3 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M25631; AAA36741.1; -
DR EMBL; X04665; CAA28370.1; -
DR EMBL; X14787; CAA32889.1; -
DR EMBL; J04835; AAA61178.1; -
DR EMBL; M99425; AAB59366.1; -
DR PIR; A05172; A05172.
DR PIR; A25812; A25812.
DR PIR; A26155; A26155.
DR PIR; A30140; A30140.
DR PIR; A34274; A34274.
DR GlycoSuiteDB; P07996; -
DR Genew; HGNC:11785; THBS1.
DR MIM; 188060; -
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR003129; TSPN.
DR InterPro; IPR001007; VWFC.
DR InterPro; IPR003367; tsp_3.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; tsp_1; 3.
DR Pfam; PF00093; vwc; 1.
DR Pfam; PF02210; TSPN; 1.
DR Pfam; PF02412; tsp_3; 8.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS01208; VWFC; 1. ~
KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170 THROMBOSPONDIN 1.
FT DOMAIN 19 232 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 316 373 VWFC.
FT DOMAIN 379 430 TSP TYPE-1 1.
FT DOMAIN 435 491 TSP TYPE-1 2.
FT DOMAIN 492 548 TSP TYPE-1 3.
FT DOMAIN 549 587 EGF-LIKE 1.
FT DOMAIN 588 645 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 646 690 EGF-LIKE 3.
FT DOMAIN 723 758 TSP TYPE-3 1.
FT DOMAIN 759 781 TSP TYPE-3 2.


```

FT DOMAIN 782 817 TSP TYPE-3 3.
FT DOMAIN 818 840 TSP TYPE-3 4.
FT DOMAIN 841 878 TSP TYPE-3 5.
FT DOMAIN 879 914 TSP TYPE-3 6.
FT DOMAIN 915 950 TSP TYPE-3 7.
FT DOMAIN 951 1170 C-TERMINAL.
FT SITE 926 928 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 270 270 INTERCHAIN (PROBABLE).
FT DISULFID 274 274 INTERCHAIN (PROBABLE).
FT DISULFID 551 562 BY SIMILARITY.
FT DISULFID 556 572 BY SIMILARITY.
FT DISULFID 575 586 BY SIMILARITY.
FT DISULFID 592 608 BY SIMILARITY.
FT DISULFID 599 617 BY SIMILARITY.
FT DISULFID 620 644 BY SIMILARITY.
FT DISULFID 650 663 BY SIMILARITY.
FT DISULFID 657 676 BY SIMILARITY.
FT DISULFID 678 689 BY SIMILARITY.
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 385 385 C-LINKED (MAN).
FT 394 394 /FTID=CAR_000205.
FT 394 394 O-LINKED (FUC. . .).
FT 438 438 /FTID=CAR_000206.
FT 441 441 C-LINKED (MAN).
FT 441 441 /FTID=CAR_000207.
FT 450 450 C-LINKED (MAN).
FT 450 450 /FTID=CAR_000208.
FT 498 498 O-LINKED (FUC. . .).
FT 498 498 /FTID=CAR_000209.
FT 507 507 C-LINKED (MAN).
FT 507 507 /FTID=CAR_000210.
FT 708 708 O-LINKED (FUC. . .).
FT 708 708 /FTID=CAR_000211.
FT CARBOHYD 1067 1067 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 84 84 T -> A (IN REF. 2, 3 AND 4).
FT CONFLICT 523 523 T -> A (IN REF. 2).
SQ SEQUENCE 1170 AA; 129412 MW; 69B3EDE5AE3A395E CRC64;

Query Match 20.1%; Score 28; DB 1; Length 1170;
Best Local Similarity 30.8%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
Db 385 WSEWTSCSTSCGN 397

RESULT 49
ID TOP2_ASFB7 STANDARD; PRT; 1192 AA.
AC Q00942;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA topoisomerase II (EC 5.99.1.3).
GN TOP OR P1192R.
OS African swine fever virus (strain BA71V) (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10498;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92263807; PubMed=1316688;
RA Garcia-Beato R., Freije J.M., Lopez-Otin C., Blasco R., Vinuela E.;
RT "A gene homologous to topoisomerase II in African swine fever virus.";
RL Virology 188:938-947(1992).
RN [2]
RP COMPLETE GENOME.
RA Yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C.,
RA Rodriguez J.F., Vinuela E.;
RT "Analysis of the complete nucleotide sequence of African swine fever
virus.";
```

```

RL Virology 208:249-278(1995).
CC -!- FUNCTION: CAN INTRODUCE NEGATIVE SUPERHELICAL TURNS INTO
CC DOUBLE-STRANDED CIRCULAR DNA.
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -!- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M88699; AAA42735.1; -.
DR EMBL; U18466; AAA65341.1; -.
DR PIR; B42549; ISXFAS.
DR HSSP; P06786; 1BGW.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR001241; DNA_topoisoiI.
DR InterPro; IPR002205; DNA_topoisoiV.
DR Pfam; PF00204; DNA_gyraseB; 1.
DR Pfam; PF00521; DNA_topoisoiV; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR PRINTS; PR00418; TPI2FAMILY.
DR ProDom; PD000742; DNA_topoisoiV; 1.
DR SMART; SM00433; TOP2c; 1.
DR SMART; SM00434; TOP4c; 1.
DR PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW Topoisomerase; Isomerase; DNA-binding; ATP-binding.
FT NP_BIND 142 147 ATP (POTENTIAL).
FT ACT_SITE 800 800 DNA CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 1192 AA; 135543 MW; 13887DB36D26C667 CRC64;

Query Match 20.1%; Score 28; DB 1; Length 1192;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXQTXLXN 51
Db 599 WAKKQTSLVN 608

RESULT 50
CCMC_ECOLI STANDARD; PRT; 245 AA.
AC P33929;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heme exporter protein C (Cytochrome c-type biogenesis protein ccmC).
GN CCMC OR B2199 OR Z3456 OR ECS3088.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / BHB2600;
RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
RA Church G.M.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS, CHARACTERIZATION, AND GENE NAME.
RX MEDLINE=95362656; PubMed=7635817;
RA Thoeny-Meyer L., Fischer F., Kunzler P., Ritz D., Hennecke H.;
RT "Escherichia coli genes required for cytochrome c maturation.";
RL J. Bacteriol. 177:4321-4326(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
```


RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: REQUIRED FOR THE EXPORT OF HEME TO THE PERIPLASM FOR THE
CC BIOGENESIS OF C-TYPE CYTOCHROMES.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- SIMILARITY: BELONGS TO THE CCMC/CYCZ/HELC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00008; AAA16391.1; ALT_SEQ.
DR EMBL; AE000309; AAC75259.1; -.
DR EMBL; AE005452; AAG57334.1; -.
DR EMBL; AP002560; BAB36511.1; -.
DR EcoGene; EG12057; ccmC.
DR InterPro; IPR002541; CytC_asm.
DR InterPro; IPR003557; CytC_biog_CcmC.
DR Pfam; PF01578; CytC_asm; 1.
DR PRINTS; PR01386; CCMCBIOGNIS.
DR TIGRFAMS; TIGR01191; ccmC; 1.
KW Cytochrome c-type biogenesis; Transport; Transmembrane;
KW inner membrane; Complete proteome.
FT TRANSMEM 21 41 POTENTIAL.
FT TRANSMEM 64 84 POTENTIAL.
FT TRANSMEM 91 111 POTENTIAL.
FT TRANSMEM 129 149 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
FT TRANSMEM 206 226 POTENTIAL.
SQ SEQUENCE 245 AA; 27885 MW; A98F2F754948EA60 CRC64;

Query Match 19.4%; Score 27; DB 1; Length 245;
Best Local Similarity 29.4%; Pred. No. 57;
Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

OY 39 WXXWXXXQTXLNEXXL 55
| | | : | |
Db 119 WGTWVWDARLTSELVL 135

Search completed: July 18, 2003, 15:59:27
Job time : 24 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:51:48 ; Search time 81 Seconds
(without alignments)
165.346 Million cell updates/sec

Title: US-09-955-502-1
Perfect score: 139
Sequence: 1 MXXRXXCXXXXXXX.....QTXLXNEXLXXXXXXRXX 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_21:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	49	35.3	90	16	Q8ZHE7
2	48	34.5	91	16	Q8XFF6
3	47	33.8	90	16	Q9KUR4
4	47	33.8	90	16	Q9CLB9
5	46	33.1	88	16	Q9JQP5
6	45	32.4	91	16	Q8Y010
7	44	31.7	90	16	Q9HU36
8	44	31.7	105	16	Q9PC73
9	35	25.2	1767	5	Q9W3P1
10	34	24.5	221	5	Q8SS09
11	34	24.5	764	4	Q9HAP8
12	33	23.7	132	2	Q9EY23
13	33	23.7	159	2	Q93RG2
14	33	23.7	385	5	Q9VJP0
15	33	23.7	393	16	Q9RJX2
16	33	23.7	444	5	Q9NK86

17	33	23.7	1235	16	Q8Y511	Q8y511 listeria mo
18	32	23.0	380	12	Q9YW02	Q9yw02 melanoplus
19	32	23.0	420	16	P73615	P73615 synechocyst
20	32	23.0	518	5	O97352	O97352 trypanosoma
21	32	23.0	1194	4	Q9P2F6	Q9p2f6 homo sapien
22	31	22.3	493	2	O30495	O30495 pseudomonas
23	31	22.3	598	16	Q8XYX5	Q8xyx5 ralstonia s
24	31	22.3	652	2	Q93TK2	Q93tk2 streptococc
25	31	22.3	751	10	Q9AY55	Q9ay55 oryza sativ
26	30	21.6	99	16	Q8XGRI	Q8xgr1 salmonella
27	30	21.6	99	16	Q8XAH2	Q8xah2 escherichia
28	30	21.6	124	13	Q9I8Q8	Q9i8q8 gallus gall
29	30	21.6	189	11	Q9D5B7	Q9d5b7 mus musculu
30	30	21.6	235	5	Q8T161	Q8t161 dictyosteli
31	30	21.6	237	5	Q8T623	Q8t623 dictyosteli
32	30	21.6	287	10	Q9ZRT9	Q9zrt9 arabidopsi
33	30	21.6	331	10	Q9SVU2	Q9svu2 arabidopsi
34	30	21.6	347	10	Q94AF3	Q94af3 arabidopsi
35	30	21.6	385	10	O80521	O80521 arabidops
36	30	21.6	532	10	Q93YW4	Q93yw4 arabidopsi
37	30	21.6	536	10	Q9LPP3	Q9lpp3 arabidopsi
38	30	21.6	548	10	Q9CAF0	Q9caf0 arabidopsi
39	30	21.6	711	16	P96855	P96855 mycobacteri
40	30	21.6	744	4	Q9P2E0	Q9p2e0 homo sapien
41	30	21.6	767	16	Q8XRH3	Q8xrh3 ralstonia s
42	30	21.6	823	5	Q23064	Q23064 caenorhabdi
43	30	21.6	915	10	Q949G7	Q949g7 malus flori
44	30	21.6	921	4	Q9Y4C2	Q9y4c2 homo sapien
45	30	21.6	980	10	Q949G8	Q949g8 malus flori

ALIGNMENTS

RESULT 1
Q8ZHE7
ID Q8ZHE7 PRELIMINARY; PRT; 90 AA.
AC Q8ZHE7;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hypothetical protein YPO0953.
GN YPO0953.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
DR EMBL; AJ414145; CAC89796.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 90 AA; 10707 MW; C7374E6855653F65 CRC64;

Query Match 35.3%; Score 49; DB 16; Length 90;
Best Local Similarity 23.6%; Pred. No. 0.00087;
Matches 13; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 MXXRXXCXXXXXXX.....QTXLXNEXLXXXXXXRXX 55
Db 1 MSRTIFCTFLKKDAERQDFQLYPGEIGKRIYNEISKEAWSQWTKOTMLINEKKL 55

RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montei-ro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159(2000).
DR EMBL; AE04010; AAF84714.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 105 AA; 12200 MW; CE751972B8399873 CRC64;

Query Match 31.7%; Score 44; DB 16; Length 105;
Best Local Similarity 47.1%; Pred. No. 0.014; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 9;

QY 39 WXXWXXXQTXLXNEXL 55
Db 54 WAAWLHVQTMLINENRL 70

RESULT 9
Q9W3P1 PRELIMINARY; PRT; 1767 AA.
AC Q9W3P1;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE CG15327 protein.
GN CG15327.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balilew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AE003442; AAF46280.1; -;
DR FlyBase; FBgn0029967; CG15327.
DR InterPro; IPR001311; SBP/glu_receptor.
SQ SEQUENCE 1767 AA; 202884 MW; 243276182343EEC6 CRC64;

Query Match 25.2%; Score 35; DB 5; Length 1767;
Best Local Similarity 35.7%; Pred. No. 25;
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
Db 1183 WSRWELSESYL RNE 1196

RESULT 10
Q8SS09 PRELIMINARY; PRT; 221 AA.
AC Q8SS09;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ADP ribosylation factor-like GTP binding protein.
GN ECU05_0090.
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomar F.,
RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi";
RL Nature 414:450-453(2001).
DR EMBL; AL590445; CAD26526.1; -;
SQ SEQUENCE 221 AA; 25102 MW; 6BA2DEA6BE00AB98 CRC64;

Query Match 24.5%; Score 34; DB 5; Length 221;
Best Local Similarity 28.6%; Pred. No. 5.8;
Matches 4; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
Db 160 WKSLSQETGIENQ 173

RESULT 11
Q9HAP8 PRELIMINARY; PRT; 764 AA.
ID Q9HAP8
AC Q9HAP8;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Testis development protein PRTD.
OS Homo sapiens (Human).

Query Match 23.7%; Score 33; DB 5; Length 385;
Best Local Similarity 41.2%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| : || || |
Db 199 WQHFYELQTTLYNEFL 215

RESULT 15

Q9RJX2 PRELIMINARY; PRT; 393 AA.
AC Q9RJX2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Putative acyl-CoA dehydrogenase.
GN SC01198 OR SCG111A.29C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmid and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; ALI33210; CAB61610.1; -
DR InterPro; IPR001552; Acyl-CoA_dh.
DR InterPro; IPR001525; C5_DNA_meth.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00095; C5_MTASE_2; UNKNOWN_1.
SQ SEQUENCE 393 AA; 43350 MW; E3E7D6B714C6BF8F CRC64;

Query Match 23.7%; Score 33; DB 16; Length 393;
Best Local Similarity 54.5%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNE 52
| : || || |
Db 220 WRVACTTLNNE 230

RESULT 16

Q9NK86 PRELIMINARY; PRT; 444 AA.
AC Q9NK86;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Hypothetical 51.9 kDa protein.
GN BG:DS04862.2 OR CG15256.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=99403001; PubMed=10471707;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA Celniker S., Rubin G.M.;
RT "An exploration of the sequence of a 2.9-Mb region of the genome of
Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Celnikier S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003413; AAF4947.1; -
DR FlyBase; FBgn0028880; BG:DS04862.2.
KW Hypothetical protein.
SQ SEQUENCE 444 AA; 51889 MW; BAFE69A2E507B82C CRC64;

Query Match 23.7%; Score 33; DB 5; Length 444;
Best Local Similarity 41.2%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| : || || |
Db 258 WQHFYELQTTLYNEFL 274

RESULT 17

Q8Y511 PRELIMINARY; PRT; 1235 AA.
AC Q8Y511;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical protein lmo2267.
GN LMO2267.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-E / SEROVAR 1/2A;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of *Listeria species*.";
RL Science 294:849-852(2001).
DR EMBL; AL591982; CAD00345.1; -.
DR ListiList; LMO02267; -.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR000212; UvrD-helicase.
DR Pfam; PF00580; UvrD-helicase; 1.
DR PROSITE; PS00225; CRYSTALLIN_BETACAMMA; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1235 AA; 142655 MW; E0329AD6F07E4716 CRC64;

Query Match 23.7%; Score 33; DB 16; Length 1235;
Best Local Similarity 30.8%; Pred. No. 51;
Matches 4; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
Db 276 WSSWTHLKTSIEN 288

RESULT 18
Q9YW02 PRELIMINARY; PRT; 380 AA.
AC Q9YW02;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF MSV090 putative Molluscum contagiosum virus MC121L (vaccinia A16L)
DE homolog, similar to GB:U60315.
GN MSV090.
OS Melanoplus sanguinipes entomopoxvirus (MSEPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OC Entomopoxvirus B.
OX NCBI_TaxID=83191;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUCSON;
RX MEDLINE=99102612; PubMed=9847359;
RA Afonso C.L., Tulman E.R., Lu Z., Oma E., Kutish G.F., Rock D.L.;
RT "The genome of Melanoplus sanguinipes entomopoxvirus.";
RL J. Virol. 73:533-552(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TUCSON;
RA Afonso C.L., Tulman E.R., Lu Z., Oma E., Kutish G.F., Rock D.L.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF063866; AAC97640.1; -.
DR InterPro; IPR004251; DUF230.
DR Pfam; PF03003; DUF230; 1.
SQ SEQUENCE 380 AA; 42842 MW; FF954C5E6587C92D CRC64;

Query Match 23.0%; Score 32; DB 12; Length 380;
Best Local Similarity 42.9%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 42 WXXXTXLXNEXXL 55
Db 186 WLESQTKLNNDIAL 199

RESULT 19
P73615
ID P73615 PRELIMINARY; PRT; 420 AA.
AC P73615;

DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein slr1865.
GN SLR1865.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90908; BAA17660.1; -.
DR InterPro; IPR000567; SBP_bac_1.
DR Pfam; PF01547; SBP_bacterial_1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 420 AA; 47851 MW; D617A38D9E8E958C CRC64;

Query Match 23.0%; Score 32; DB 16; Length 420;
Best Local Similarity 35.7%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
Db 150 WRFWEVAQTNLIKQ 163

RESULT 20
O97352 PRELIMINARY; PRT; 518 AA.
ID O97352
AC O97352;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ILTat 1.6l metacyclic VSG protein.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20017521; PubMed=10551359;
RA Graham S.V., Terry S., Barry J.D.;
RT "A structural and transcription pattern for variant surface
RT glycoprotein gene expression sites used in metacyclic stage
RT Trypanosoma brucei.";
RL Mol. Biochem. Parasitol. 103:141-154(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91081307; PubMed=2175429;
RA Matthews K.R., Shiels P.G., Graham S.V., Cowan C., Barry J.D.;
RT "Duplicative activation mechanisms of two trypanosome telomeric VSG
RT genes with structurally simple 5' flanks.";
RL Nucleic Acids Res. 18:7219-7227(1990).
DR EMBL; AJ012199; CAA09956.1; -.
SQ SEQUENCE 518 AA; 55736 MW; 800D002074229468 CRC64;

Query Match 23.0%; Score 32; DB 5; Length 518;
Best Local Similarity 35.3%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
Db 108 WEEWKNSATRLKEAVL 124

RESULT 21
Q9P2F6 PRELIMINARY; PRT; 1194 AA.
AC Q9P2F6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE KIAA1391 protein (Fragment).
GN KIAA1391.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
DR EMBL; AB037812; BAA92629.1; -.
DR HSSP; Q07960; IRGP.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000159; RA_domain.
DR InterPro; IPR000198; RhOGAP.
DR Pfam; PF00620; RhOGAP; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00324; RhOGAP; 1.
FT NON_TER 1
SQ SEQUENCE 1194 AA; 133015 MW; C4BB7B71F4CC9F59 CRC64;

Query Match 23.0%; Score 32; DB 4; Length 1194;
Best Local Similarity 38.5%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
Db 340 WAFWRGSSTHLDN 352

RESULT 22
Q30495 PRELIMINARY; PRT; 493 AA.
AC Q30495;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Xylulose kinase (EC 2.7.1.17) (Xylulokinase).
GN MTLY OR XYL6.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN=DSM 50106;
RX MEDLINE=98121321; PubMed=9461423;
RA Bruenker P., Altenbuchner J., Mattes R.;
RT "Structure and function of the genes involved in mannitol, arabitol
RT and glucitol utilization from Pseudomonas fluorescens DSM50106."
RL Gene 206:117-126(1998).
CC -!- CATALYTIC ACTIVITY: ATP + D-XYLULOSE = ADP + D-XYLULOSE 5-
CC PHOSPHATE.
CC -!- INDUCTION: BY MANNITOL, ARABITOL AND SORBITOL. NOT INDUCED BY
CC FRUCTOSE OR XYLULOSE.
CC -!- MISCELLANEOUS: THIS ENZYME IS SPECIFIC FOR XYLULOSE.
CC -!- MISCELLANEOUS: THE MTLY PROTEIN IS ENCODED BY THE MTL EFGKDYZ
CC OPERON. THIS OPERON ENCODES PROTEINS FOR THE UPTAKE AND
CC UTILIZATION OF MANNITOL, ARABITOL AND SORBITOL.
CC -!- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCONOKINASE /
CC GLYCEROKINASE / XYLULOKINASE FAMILY.

DR EMBL; AF007800; AAC04473.1; -.
DR InterPro; IPR000577; FGGY_kin.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00370; FGGY; 1.
DR Pfam; PF02782; FGGY_C; 1.
DR TIGRFAMS; TIGR01312; XylB; 1.
DR PROSITE; PS00933; FGGY_KINASES_1; FALSE_NEG.
DR PROSITE; PS00445; FGGY_KINASES_2; UNKNOWN_1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
KW Transferase; Kinase; Xylose metabolism.
SQ SEQUENCE 493 AA; 52057 MW; 6E66792F8E5B5C65 CRC64;

Query Match 22.3%; Score 31; DB 2; Length 493;
Best Local Similarity 42.9%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNEXXL 55
Db 104 WCDTETALENERLL 117

RESULT 23
Q8YX5 PRELIMINARY; PRT; 598 AA.
AC Q8YX5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable poly-beta-hydroxybutyrate polymerase transmembrane protein
DE (EC 2.3.1.-).
GN PHBC OR RSC1631 OR RS03994.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL; AL646065; CAD15333.1; -.
DR InterPro; IPR000073; Abhydrolase.
DR Pfam; PF00561; abhydrolase; 1.
KW Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 598 AA; 65607 MW; B50BC3E94AE8D98E CRC64;

Query Match 22.3%; Score 31; DB 16; Length 598;
Best Local Similarity 35.7%; Pred. No. 74;
Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
Db 445 WYCWLRLRHTYLQND 458

RESULT 24
Q93TK2 PRELIMINARY; PRT; 652 AA.
AC Q93TK2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Amylase-binding protein B.
GN ABPB.
OS Streptococcus gordonii.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;

OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID=1302;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Li L.N., Scannapieco F.A.;
 RT "Identification and Analysis of the Amylase-Binding Protein B and Gene
 from Streptococcus gordonii.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF354648; AAK52749.1; -
 DR InterPro; IPR005322; Peptidase_U34.
 DR Pfam; PF03577; Peptidase_U34; 1.
 SQ SEQUENCE 652 AA; 72741 MW; C4092EFFE53DFD4F CRC64;
 Query Match 22.3%; Score 31; DB 2; Length 652;
 Best Local Similarity 28.6%; Pred. No. 80;
 Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 OY 39 WXXWXXXQTXLXNE 52
 | | | | |
 Db 422 WTVWHIDQMAIKNQ 435
 RESULT 25
 Q9AY55 PRELIMINARY; PRT; 751 AA.
 ID Q9AY55
 AC Q9AY55;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Putative arm repeat protein.
 GN OSJNBA0027P10.11
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Hsiao J.,
 RA Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E., Craven B.,
 RA Khalak H., Feldblyum T.V., Quackenbush J., White O., Salzberg S.L.,
 RA Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJNBA0027P10 genomic sequence.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC084763; AAG60190.1; -
 DR HSSP; P35222; 1G3J.
 DR InterPro; IPR000225; Armadillo.
 DR InterPro; IPR001810; F-box.
 DR InterPro; IPR003592; LRR_out.
 DR Pfam; PF00514; Armadillo_seg; 7.
 DR Pfam; PF00646; F-box; 2.
 DR SMART; SM00185; ARM; 6.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00370; LRR; 3.
 DR PROSITE; PS50181; FBOX; 1.
 SQ SEQUENCE 751 AA; 79515 MW; 2E33B5FEE4A26685 CRC64;
 Query Match 22.3%; Score 31; DB 10; Length 751;
 Best Local Similarity 45.5%; Pred. No. 92;
 Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 OY 39 WXXWXXXQTXL 49
 | | | | |
 Db 375 WLEWILSQTL 385
 RESULT 26
 Q8XGR1 PRELIMINARY; PRT; 99 AA.
 ID Q8XGR1
 AC Q8XGR1;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Putative inner membrane protein (Hypothetical protein STY3411).
 GN YQJK OR STM3231 OR STY3411.
 OS Salmonella typhimurium, and
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602, 601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 DR EMBL; AE008848; AAL22104.1; -
 DR EMBL; AL627278; CAD07755.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 99 AA; 11548 MW; 26F4194AE256CE35 CRC64;
 Query Match 21.6%; Score 30; DB 16; Length 99;
 Best Local Similarity 28.6%; Pred. No. 22;
 Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 OY 39 WXXWXXXQTXLXNE 52
 | | | | |
 Db 82 WSAWRLVKTTLRQQ 95
 RESULT 27
 Q8XAH2 PRELIMINARY; PRT; 99 AA.
 ID Q8XAH2
 AC Q8XAH2;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Orf, hypothetical protein.
 GN Z4454 OR ECS3982.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";

```
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
DR EMBL; AE005539; AAG58233.1; -.
DR EMBL; AP002564; BAB37405.1; -.
KW Complete proteome.
SQ SEQUENCE 99 AA; 11781 MW; C6B9170493DF4086 CRC64;

Query Match 21.6%; Score 30; DB 16; Length 99;
Best Local Similarity 28.6%; Pred. No. 22;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | | :
Db 82 WSAWRLVKTTLKQQ 95

RESULT 28
Q9I8Q8 PRELIMINARY; PRT; 124 AA.
AC Q9I8Q8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Bone morphogenetic protein 7 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20307867; PubMed=10848589;
RA Monroe D.G., Jin D.F., Sanders M.M.;
RT "Estrogen Opposes the Apoptotic Effects of Bone Morphogenetic Protein
RT 7 on Tissue Remodeling.";
RL Mol. Cell. Biol. 20:4626-4634(2000).
DR EMBL; AF223970; AAF34758.1; -.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00688; TGFb_propeptide; 1.
FT NON_TER 124 124
SQ SEQUENCE 124 AA; 13794 MW; F70104D9196A4248 CRC64;

Query Match 21.6%; Score 30; DB 13; Length 124;
Best Local Similarity 35.7%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | |
Db 23 WLRWVLADFTILDNE 36

RESULT 29
Q9D5B7 PRELIMINARY; PRT; 189 AA.
AC Q9D5B7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 4930469G21Rik protein.
GN 4930469G21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK015528; BAB29884.1; -.
DR MGD; MGI:1922183; 4930469G21Rik.
SQ SEQUENCE 189 AA; 21841 MW; C445DF77CDEE361B CRC64;

Query Match 21.6%; Score 30; DB 11; Length 189;
Best Local Similarity 29.4%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNEXXL 55
| | : | | :
Db 99 WMKWTRHQKLLKKQASL 115

RESULT 30
Q8T161 PRELIMINARY; PRT; 235 AA.
AC Q8T161;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Protein-S isoprenylcysteine O-methyltransferase.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC117070; AAM09306.1; -.
KW Transferase; Methyltransferase.
SQ SEQUENCE 235 AA; 27177 MW; CCDCAA6C8AF85014 CRC64;

Query Match 21.6%; Score 30; DB 5; Length 235;
Best Local Similarity 28.6%; Pred. No. 51;
Matches 4; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : | |
Db 195 WASWSFFSQRIENE 208

RESULT 31
Q8T623 PRELIMINARY; PRT; 237 AA.
ID Q8T623
```

AC Q8T623;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Prenyl cysteine carboxyl methyltransferase.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Y., Stock J.B., Cox E.C.;
RT "Prenyl Cysteine Carboxyl Methyltransferase in Dictyostelium.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF487784; AAL99548.1; -;
KW Transferase; Methyltransferase.
SQ SEQUENCE 237 AA; 27379 MW; 3F6961685B01C5DE CRC64;

Query Match 21.6%; Score 30; DB 5; Length 237;
Best Local Similarity 28.6%; Pred. No. 51;
Matches 4; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : ||
Db 197 WASWSFFSQRIENE 210

RESULT 32
Q9ZRT9
ID Q9ZRT9 PRELIMINARY; PRT; 287 AA.
AC Q9ZRT9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE LEA-like protein.
GN LEA32.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA; TISSUE=DRY SEEDS;
RC Bianchi M.W., Javot H., Vartanian N.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=COLUMBIA;
RC Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
DR EMBL; AJ131342; CAA10352.1; -;
DR EMBL; AB022219; BAB02041.1; -;
DR InterPro; IPR004238; LEA.
DR Pfam; PF02987; LEA; 3.
SQ SEQUENCE 287 AA; 31312 MW; C744A7B1267BE827 CRC64;

Query Match 21.6%; Score 30; DB 10; Length 287;
Best Local Similarity 28.6%; Pred. No. 61;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : ||
Db 143 WTCWAKEKIGIKNE 156

RESULT 33
Q9SVU2
ID Q9SVU2 PRELIMINARY; PRT; 331 AA.
AC Q9SVU2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Hypothetical 37.2 kDa protein.
GN F16A16.150 OR AF4G28740.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Brandt P., Dose S., Jarke D., Scharfe M., Schon O.,
RA Hoheisel J., Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Brandt P., Dose S., Jarke D., Scharfe M., Schon O., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035353; CAA22978.1; -;
DR EMBL; AL161573; CAB81462.1; -;
KW Hypothetical protein.
SQ SEQUENCE 331 AA; 37227 MW; B37B39CA026E3CD9 CRC64;

Query Match 21.6%; Score 30; DB 10; Length 331;
Best Local Similarity 38.5%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | : ||
Db 267 WEKWLNEQKKLAN 279

RESULT 34
Q94AF3
ID Q94AF3 PRELIMINARY; PRT; 347 AA.
AC Q94AF3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE AT4g28740/F16A16_150.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Koesema E., Chen H., Cheuk R., Kim C.J., Meyers M.C., Shinn P.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Gibson H.A.,
RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J.,
RA Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]

RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Bowser L., Carninci P., Jones T., Kamiya A., Karlin-Neumann G.,
RA Hayashizaki Y., Ishida J., Lin J., Miranda M., Narusaka M.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.,
RT "Arabidopsis ORF clones";
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY048223; AAK82486.1; -;
DR EMBL; AY091705; AAM10304.1; -;
SQ SEQUENCE 347 AA; 38989 MW; B9277377859D23E8 CRC64;

Query Match 21.6%; Score 30; DB 10; Length 347;
Best Local Similarity 38.5%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
Db 283 WEKWLNEQKKLAN 295

RESULT 35
O80521
ID O80521 PRELIMINARY; PRT; 385 AA.
AC O80521;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F14J9.4 protein.
GN F14J9.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Federspiel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
RA Au M., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,
RA Oji O., Osborne B.I., Shinn P., Sun H., Toriumi M., Vysotskaia V.S.,
RA Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AC003970; AAC33198.1; -;
DR InterPro; IPR000620; DUF6.
DR Pfam; PF00892; DUF6; 2.
SQ SEQUENCE 385 AA; 42125 MW; F3A6266C889633B2 CRC64;

Query Match 21.6%; Score 30; DB 10; Length 385;
Best Local Similarity 36.4%; Pred. No. 82;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
Db 208 WAAWFIIQTKM 218

RESULT 36
Q93YW4
ID Q93YW4 PRELIMINARY; PRT; 532 AA.
AC Q93YW4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical 60.5 kDa protein.
GN At5G38600.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Full Length cDNA of gene At5g38600 (GI:15240959).";
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY059738; AAL24150.1; -;
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00098; zf-CCHC; 1.
KW Hypothetical protein.
SQ SEQUENCE 532 AA; 60549 MW; B8BF45F18D8BD8F2 CRC64;

Query Match 21.6%; Score 30; DB 10; Length 532;
Best Local Similarity 28.6%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
Db 146 WSEWEAENTSLAQD 159

RESULT 37
Q9LPP3
ID Q9LPP3 PRELIMINARY; PRT; 536 AA.
AC Q9LPP3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE F18K10.11 protein.
GN F18K10.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F18K10 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: CONTAINS 2 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; AC013428; AAF76355.1; -;
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 1.
DR SMART; SM00320; WD40; 1.
DR PROSITE; PS50082; WD_REPEATS_2; 1.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 536 AA; 60261 MW; 9FFA2546ADF185D5 CRC64;

Query Match 21.6%; Score 30; DB 10; Length 536;
Best Local Similarity 45.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXWXXXQTXLXNE 52
Db 91 WRVKQTDIANE 101

RESULT 38
Q9CAF0
ID Q9CAF0 PRELIMINARY; PRT; 548 AA.
AC Q9CAF0;


```
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 61.7 kDa protein.
GN Fl3M14.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseld M.,
RA Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA De Simone V., Choisme N., Artiguenave F., Robert C., Brottier P.,
RA Wincker P., Cattolico L., Weissbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,
RA Reichelt J., Scharfe M., Schoen O., Bagues M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujil C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:820-822(2000).
CC -!- SIMILARITY: CONTAINS 2 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; AC011560; AAG51373.1; -.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 1.
DR SMART; SM00320; WD40; 4.
DR PROSITE; PS50082; WD_REPEATS_2; 1.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 548 AA; 61692 MW; 7E3DAA998DF78377 CRC64;

Query Match 21.6%; Score 30; DB 10; Length 548;
Best Local Similarity 45.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNE 52
| | | |
Db 103 WRVKQTDIANE 113

RESULT 39
P96855
ID P96855 PRELIMINARY; PRT; 711 AA.
AC P96855;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 74.5 kDa protein (Acyl-CoA dehydrogenase, putative).
GN FADE34 OR RV3573C OR MT3678 OR MTCY06G11.20C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
```

```
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z92774; CAB07147.1; -.
DR EMBL; AE007169; AAK48037.1; -.
DR TIGR; MT3678; -.
DR Tuberculist; RV3573C; -.
DR InterPro; IPR001552; Acyl-CoA_dh.
DR Pfam; PF00441; Acyl-CoA_dh; 2.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 557 557 F -> I (IN REF. 2).
FT CONFLICT 630 630 R -> W (IN REF. 2).
SQ SEQUENCE 711 AA; 74577 MW; 52B63290EFE196D CRC64;

Query Match 21.6%; Score 30; DB 16; Length 711;
Best Local Similarity 45.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNE 52
| | | |
Db 571 WRLARTTLANE 581

RESULT 40
Q9P2E0
ID Q9P2E0 PRELIMINARY; PRT; 744 AA.
AC Q9P2E0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE KIAA1407 protein (Fragment).
GN KIAA1407.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirotsawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:65-73(2000).
DR -- EMBL; AB037828; BAA92645.1; -.
```


FT NON_TER 1 1
SQ SEQUENCE 744 AA; 88339 MW; 1C232B0C3022FCCE CRC64;

Query Match 21.6%; Score 30; DB 4; Length 744;
Best Local Similarity 45.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNE 52
| : | | | |
Db 734 WLSKTSLVNE 744

RESULT 41
Q8XRH3 PRELIMINARY; PRT; 767 AA.
AC Q8XRH3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative transmembrane protein.
GN RSP0881 OR RS01652.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siquier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646081; CAD18032.1; -;
DR InterPro; IPR001791; Laminin_G.
DR PROSITE; PS50025; LAM_G_DOMAIN; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 767 AA; 79358 MW; 963F5DC0E001B50E CRC64;

Query Match 21.6%; Score 30; DB 16; Length 767;
Best Local Similarity 44.4%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXXQTXQT 47
| | | | |
Db 627 WVFWTAAQT 635

RESULT 42
Q23064 PRELIMINARY; PRT; 823 AA.
AC Q23064;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 93.0 kDa protein.
GN W01A11.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";

RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Blanchard M., Bradshaw H.;
RT "The sequence of C. elegans cosmid W01A11.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U64852; AAB04966.1; -;
DR InterPro; IPR001875; DED.
DR SMART; SM00031; DED; 1.
KW Hypothetical protein.
SQ SEQUENCE 823 AA; 92955 MW; AEE22220B61D1F14 CRC64;

Query Match 21.6%; Score 30; DB 5; Length 823;
Best Local Similarity 30.8%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXXQTXLXN 51
| | | | |
Db 641 WDAWNSRQNDIRN 653

RESULT 43
Q949G7 PRELIMINARY; PRT; 915 AA.
AC Q949G7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hcrvf3 protein.
GN HCRVF3.
OS Malus floribunda.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=138912;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. FLORINA;
RA Vinatzer B.A., Patocchi A., Gianfranceschi L., Tartarini S.,
RA Zhang H.B., Gessler C., Sansavini S.;
RT "Apple (Malus sp.) contains receptor-like genes homologous to the Cf
resistance gene family of tomato with a cluster of such genes co-
segregating with Vf apple scab resistance.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ297741; CAC40827.1; -;
DR InterPro; IPR001611; LRR.
DR Pfam; PF00560; LRR; 22.
SQ SEQUENCE 915 AA; 102689 MW; 028A41C8719D39EC CRC64;

Query Match 21.6%; Score 30; DB 10; Length 915;
Best Local Similarity 45.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXXQTXL 49
| | | | |
Db 433 WPMWLRQTQL 443

RESULT 44
Q9Y4C2 PRELIMINARY; PRT; 921 AA.
AC Q9Y4C2; Q9BW63;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE KIAA0738 protein.

```
GN KIAA0738.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=99087487; PubMed=9872452;
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XI.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:277-286(1998).
RN [2]
RP SEQUENCE OF 1-919 FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB018281; BAA34458.1; -.
DR EMBL; BC000609; AAH00609.1; -.
SQ SEQUENCE 921 AA; 102139 MW; 03D39FDF90561F74 CRC64;

Query Match 21.6%; Score 30; DB 4; Length 921;
Best Local Similarity 36.4%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
Db 824 WNAWTALEYTL 834

RESULT 45
Q949G8
ID Q949G8 PRELIMINARY; PRT; 980 AA.
AC Q949G8;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hcrvf2 protein.
GN HCRVF2.
OS Malus floribunda.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=138912;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. FLORINA;
RA Vinatzer B.A., Patocchi A., Gianfranceschi L., Tartarini S.,
RA Zhang H.B., Gessler C., Sansavini S.;
RT "Apple (Malus sp.) contains receptor-like genes homologous to the Cf
RT resistance gene family of tomato with a cluster of such genes co-
RT segregating with Vf apple scab resistance.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ297740; CAC40826.1; -.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00560; LRR; 23.
SQ SEQUENCE 980 AA; 109297 MW; DD2DB07FCA3F0A72 CRC64;

Query Match 21.6%; Score 30; DB 10; Length 980;
Best Local Similarity 45.5%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
Db 498 WPMWLRQTQQL 508

RESULT 46
Q949G9
ID Q949G9 PRELIMINARY; PRT; 1015 AA.
```

```
AC Q949G9;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hcrvf1 protein.
GN HCRVF1.
OS Malus floribunda.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=138912;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. FLORINA;
RA Vinatzer B.A., Patocchi A., Gianfranceschi L., Tartarini S.,
RA Zhang H.B., Gessler C., Sansavini S.;
RT "Apple (Malus sp.) contains receptor-like genes homologous to the Cf
RT resistance gene family of tomato with a cluster of such genes co-
RT segregating with Vf apple scab resistance.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ297739; CAC40825.1; -.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00560; LRR; 24.
SQ SEQUENCE 1015 AA; 113893 MW; 73B3A971B312F293 CRC64;

Query Match 21.6%; Score 30; DB 10; Length 1015;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
Db 552 WPMWLRQTQQL 562

RESULT 47
Q95WB6
ID Q95WB6 PRELIMINARY; PRT; 1041 AA.
AC Q95WB6;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE UNC-83.
GN UNC-83.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Starr D.A., Hermann G., Malone C., Priess J., Horvitz H.R., Han M.;
RT "UNC-83, a novel component of the nuclear envelope is essential for
RT nuclear migration.";
RL Development 0:0-0(2001).
DR EMBL; AF338767; AAL15621.1; -.
DR InterPro; IPR003660; HAMP.
DR SMART; SM00304; HAMP; 1.
SQ SEQUENCE 1041 AA; 117821 MW; BB403A9A7C41A01F CRC64;

Query Match 21.6%; Score 30; DB 5; Length 1041;
Best Local Similarity 30.8%; Pred. No. 2.1e+02;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
Db 859 WDAWNSRQNDIRN 871

RESULT 48
O49957
ID O49957 PRELIMINARY; PRT; 402 AA.
AC O49957;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
```

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Ribulose biphosphate carboxylase large chain (EC 4.1.1.39) (RuBisCO
DE large subunit) (Fragment).
GN RBCL.
OS Trichomanes diaphanum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Filicophyta; Filicopsida; Filicales; Hymenophyllaceae; Trichomanes.
OX NCBI_TaxID=58827;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISEM H3301;
RA Dubuisson J.Y.;
RL Thesis (1996), University of Montpellier II, Montpellier, France.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ISEM H3301;
RX MEDLINE=97446271; PubMed=9299219;
RA Dubuisson J.Y.;
RT "rbcl sequences: a promising tool for the molecular systematics of the
RT fern genus Trichomanes (Hymenophyllaceae).";
RL Mol. Phylogenet. Evol. 8:128-138(1997).
CC -!- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF D-
CC RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) = 3-
CC PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) = 2 3-
CC PHOSPHO-D-GLYCERATE.
CC -!- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
DR EMBL; Y09191; CAA70390.1; -.
DR HSSP; P04718; 9RUB.
DR InterPro; IPR000685; RuBisCO_large.
DR Pfam; PF00016; RuBisCO_large; 1.
DR Pfam; PF02788; RuBisCO_large_N; 1.
DR PROSITE; PS00157; RUBISCO_LARGE; 1.
KW Carbon dioxide fixation; Chloroplast; Lyase; Monooxygenase;
KW Oxidoreductase; Photorespiration; Photosynthesis.
FT NON_TER 1
FT NON_TER 402
SQ SEQUENCE 402 AA; 44247 MW; E03B049054DE4F99 CRC64;

Query Match 21.2%; Score 29.5; DB 10; Length 402;
Best Local Similarity 23.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 23 PXXGXGXXXXXXXWXX-WXXXQTXLXN 51
Db 26 PEEAGAAVAAESSTGTTVWTDGLTSLDN 55

RESULT 49
Q9RYU9 ID Q9RYU9 PRELIMINARY; PRT; 546 AA.
AC Q9RYU9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein DRA0207.
GN DRA0207.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001863; AAF12422.1; -.
DR TIGR; DRA0207; -.
DR InterPro; IPR003790; DUF187.
DR Pfam; PF02638; DUF187; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 546 AA; 58743 MW; 5081C9F91BFD30D7 CRC64;

Query Match 21.2%; Score 29.5; DB 16; Length 546;
Best Local Similarity 38.9%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 39 WXXWXXXQ-TXLXNEXXL 55
Db 281 WSAWRRQOVNLVRETAL 298

RESULT 50
P74350 ID P74350 PRELIMINARY; PRT; 124 AA.
AC P74350;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein slr1634.
GN SLR1634.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yanada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90914; BAA18444.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 124 AA; 13258 MW; 6969F49A5485C6B1 CRC64;

Query Match 20.9%; Score 29; DB 16; Length 124;
Best Local Similarity 28.6%; Pred. No. 46;
Matches 4; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
Db 97 WLSWLHGETNLHK 110

Search completed: July 18, 2003, 15:58:58
Job time : 84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2003, 15:51:29 ; Search time 70 Seconds
(without alignments)
123.733 Million cell updates/sec

Title: US-09-955-502-1
Perfect score: 139
Sequence: 1 MXXXXCXXXXXXXXXXXXX.....QTXLXNEXLXXXXXXXXXX 65

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	25.2	1767	22 ABB70457	Drosophila melanog
2	34	24.5	589	12 AAR10681	Polyhydroxybutyrat
3	34	24.5	589	14 AAR32190	Sequence encoded b
4	34	24.5	589	22 AAE10892	Alcaligenes eutrop
5	34	24.5	672	16 AAR71325	Poly-beta-hydroxya
6	34	24.5	764	23 ABB05596	Human testis devel
7	33	23.7	385	22 ABB70346	Drosophila melanog
8	33	23.7	1213	22 ABG29914	Novel human diagno
9	33	23.7	1235	23 ABB48680	Listeria monocytog
10	31	22.3	136	22 AAO00704	Human polypeptide

11	31	22.3	147	21 AAG32774	Zea mays protein f
12	31	22.3	154	21 AAG32773	Zea mays protein f
13	31	22.3	167	21 AAG32772	Zea mays protein f
14	31	22.3	399	22 AAU49357	Propionibacterium
15	30.5	21.9	835	22 AAU49903	Propionibacterium
16	30	21.6	99	21 AAB15978	E. coli proliferat
17	30	21.6	150	22 ABB34930	Peptide #2436 enco
18	30	21.6	150	22 ABB20340	Protein #2339 enco
19	30	21.6	150	22 AAM55737	Human brain expres
20	30	21.6	150	23 ABG37643	Human peptide enco
21	30	21.6	216	23 AAM49089	Human flavoprotein
22	30	21.6	238	21 AAG59921	Arabidopsis thalia
23	30	21.6	238	21 AAG59921	Arabidopsis thalia
24	30	21.6	314	22 ABG25940	Novel human diagno
25	30	21.6	331	21 AAG57492	Arabidopsis thalia
26	30	21.6	331	21 AAG59920	Arabidopsis thalia
27	30	21.6	347	21 AAG57491	Arabidopsis thalia
28	30	21.6	347	21 AAG59919	Arabidopsis thalia
29	30	21.6	349	22 AAU24598	Human olfactory r
30	30	21.6	363	21 AAG06011	Arabidopsis thali
31	30	21.6	363	21 AAG47673	Arabidopsis thalia
32	30	21.6	369	21 AAG06010	Arabidopsis thalia
33	30	21.6	369	21 AAG47672	Arabidopsis thalia
34	30	21.6	374	21 AAG06009	Arabidopsis thalia
35	30	21.6	374	21 AAG47671	Arabidopsis thalia
36	30	21.6	382	21 AAG22294	Arabidopsis thalia
37	30	21.6	386	21 AAG22293	Arabidopsis thalia
38	30	21.6	405	21 AAG22292	Arabidopsis thalia
39	30	21.6	526	21 AAG38504	Arabidopsis thalia
40	30	21.6	536	21 AAG38503	Arabidopsis thalia
41	30	21.6	882	22 AAB93159	Human protein sequ
42	30	21.6	1730	21 AAG46753	Arabidopsis thalia
43	30	21.6	1734	21 AAG46752	Arabidopsis thalia
44	29	20.9	58	22 ABB32215	Peptide #4866 enco
45	29	20.9	58	22 ABB37467	Peptide #4973 enco

ALIGNMENTS

RESULT 1
ABB70457
ID ABB70457 standard; Protein; 1767 AA.
XX
AC ABB70457;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 38163.
XX
KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
(PEKE) PE CORP NY.
Venter JC, Adams M, Li PWD, Myers EW;
WPI; 2001-656860/75.
DR N-PSDB; ABL14560.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -
XX Disclosure; SEQ ID NO 38163; 21pp + Sequence Listing; English.
PS
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1767 AA;

Query Match 25.2%; Score 35; DB 22; Length 1767;
Best Local Similarity 35.7%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : : | | |
Db 1183 WSRWELSESYLRNE 1196

RESULT 2
AAR10681
ID AAR10681 standard; Protein; 589 AA.
XX
AC AAR10681;
XX
DT 17-APR-1991 (first entry)
XX
DE Polyhydroxybutyrate polymerase enzyme.
XX
KW Polyester biopolymers; polyhydroxybutyrate; polyhydroxy alkanooate;
KW beta-ketothiolase; acetoacetyl CoA reductase.
XX
OS Alcaligenes eutrophus.
XX
PN WO9100917-A.
XX
PD 24-JAN-1991.
XX
PF 10-JUL-1990; 90WO-US03851.
XX
PR 10-JUL-1989; 89US-0378155.
XX
PA (MASI) MASSACHUSETTS INST TECH.
XX
PI Peoples OP, Sinskey AJ;
XX
DR WPI; 1991-051341/07.
DR N-PSDB; AAQ10502.
XX
PT Construction and modification of polyester bio:polymers - by
PT introduction of poly-hydroxy-butyrate and -alkanoate genes into
PT bacteria or plants
XX
PS Disclosure; fig 4; 64pp; English.
XX
CC This Alcaligenes eutrophus polyhydroxybutyrate (PHB) polymerase enzyme
CC is essential to the biosynthesis of PHB. The use of recombinant methods
CC for producing such enzymes, required for polyester biopolymer synthesis,
CC allows for the control and modification of the synthesis process.
CC See also AAQ10499-501 and
CC AAQ10503.
XX
SQ Sequence 589 AA;

Query Match 24.5%; Score 34; DB 12; Length 589;

Best Local Similarity 42.9%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : : | | |
Db 436 WYCWLRLRHTYLQNE 449

RESULT 3
AAR32190
ID AAR32190 standard; Protein; 589 AA.
XX
AC AAR32190;
XX
DT 30-MAY-1993 (first entry)
XX
DE Sequence encoded by the PHB synthase (phbC) gene of the
DE polyhydroxybutyrate (PHB) operon.
XX
KW Operon; polyhydroxyalkanoate; polyhydroxybutyrate synthase.
XX
OS Alcaligenes eutrophus.
XX
PN WO9302187-A.
XX
PD 04-FEB-1993.
XX
PF 13-JUL-1992; 92WO-US05786.
XX
PR 19-JUL-1991; 91US-0732243.
XX
PA (UYMA-) UNIV MADISON JAMES.
PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Dennis DE, Poirier Y, Somerville CR;
XX
DR WPI; 1993-058785/07.
DR N-PSDB; AAQ36660.
XX
PT Transgenic plants producing poly:hydroxy-alkanoate polymer(s) -
PT obtd. by transformation with DNA encoding 3-ketothiolase,
PT acetoacetyl-CoA reductase and PHA synthase
XX
PS Disclosure; Fig 2; 70pp; English.
XX
CC The nucleotide sequence of the PHB operon was obtained from Janes, B.
CC Hollar, J. and Dennis, D. in Dawes, E.A. (ed.) Novel Biodegradable
CC Polymers, Kluwer Academic Publishers, 175-190 (1990). It contains
CC the genes from PHB synthase, 3-ketothiolase and acetoacetyl-CoA
CC reductase. The inventors claim a transgenic plant material contg.
CC foreign DNA encoding a peptide which exhibits 3-ketothiolase activity,
CC pref. where the DNA is an open reading from between nucleotides
CC 2696-3877 (phb A gene), 842-2611 (phb C gene) or 3952-4692 (phb B
CC gene) of the Alcaligenes eutrophus PHB operon.
XX
SQ Sequence 589 AA;

Query Match 24.5%; Score 34; DB 14; Length 589;
Best Local Similarity 42.9%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | : : | | |
Db 436 WYCWLRLRHTYLQNE 449

RESULT 4
AAE10892
ID AAE10892 standard; Protein; 589 AA.
XX
AC AAE10892;
XX
DT 18-DEC-2001 (first entry)

XX
DE Alcaligenes eutrophus polyhydroxybutyrate (PHB) synthase.
XX
KW Polyhydroxybutyrate synthase; PHB; sphingan; food product; dessert gel;
KW jelly; jam; beverage; dairy product; gelling agent; rheological modifier;
KW industrial application.
XX
OS Alcaligenes eutrophus.
XX
PN WO200164897-A2.
XX
PD 07-SEP-2001.
XX
PF 02-MAR-2001; 2001WO-US07010.
XX
PR 02-MAR-2000; 2000US-186433P.
XX
PA (KELC) CP KELCO US INC.
XX
XX Bower S, Burke E, Harding N, Patel YN, Schneider JC, Meissner D;
PI Morrison N, Bezanson R;
XX
DR WPI; 2001-589870/66.
XX
XX Mutant strain of Sphingomonas species useful for producing sphingan,
PT has a mutation in the gene encoding a protein involved in
PT polyhydroxybutyrate synthesis -
XX
PS Example 1; Page 82-84; 98pp; English.
XX
CC The invention relates to mutant strains of Sphingomonas species which
CC have a mutation in the gene encoding a protein involved in internal
CC storage polymer polyhydroxybutyrate (PHB) synthesis that allows the
CC mutant strains to produce PHB-deficient sphingans. Sphingans are capsular
CC polysaccharides secreted by bacteria of the genus Sphingomonas. The
CC invention also relates to a process for preparing clarified sphingans
CC which are useful as gelling agents in a variety of food applications
CC for improving the taste, texture, stability and appearance of food
CC products such as dessert gels, confectionery jellies, jams, dairy
CC products, beverages, films and coatings. The sphingans are also useful
CC as rheological modifier in industrial applications such as oil-field
CC drilling and cementitious systems. The present sequence is
CC Alcaligenes eutrophus polyhydroxybutyrate (PHB) synthase which is used
CC for the generation of Sphingomonas elodea phAC fragment. phAC DNA
CC fragment encodes polyhydroxybutyrate (PHB) synthase protein.
XX
SQ Sequence 589 AA;
Query Match 24.5%; Score 34; DB 22; Length 589;
Best Local Similarity 42.9%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXLXNE 52
Db 436 WYCWLRLRHTYLQNE 449
RESULT 5
AAR71325
ID AAR71325 standard; Protein; 672 AA.
XX
AC AAR71325;
XX
DT 21-OCT-1995 (first entry)
XX
DE Poly-beta-hydroxyalkanoate-synthase.
XX
KW Poly-beta-hydroxyalkanoate-synthase; transgenic plant;
KW poly-beta-hydroxyalkanoate; poly-beta-hydroxybutyrate;
KW biodegradable thermoplastic.
XX
OS Alcaligenes eutrophus.
XX

PN WO9505472-A.
XX
PD 23-FEB-1995.
XX
PF 17-AUG-1994; 94WO-US09265.
XX
PR 17-AUG-1993; 93US-0108193.
PR 06-JUN-1994; 94US-0254357.
XX
PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Nawrath C, Poirier Y, Somerville CR;
XX
DR WPI; 1995-098770/13.
DR N-PSDB; AAR71325.
XX
PT Transgenic plant material with plastid(s) contg. the enzymes for
PT synthesis of poly:hydroxy:alkanoate(s) - express
PT poly:hydroxy:butyrate and have good growth and seed formation.
XX
PS Claim 2; Page 62-64; 88pp; English.
XX
CC The poly-beta-hydroxyalkanoate-synthase gene (phbC) from A.
CC eutrophus is cloned under the control of an Arabidopsis thaliana
CC seed storage protein promoter for plastid tissue-specific
CC gene expression in a transgenic plant. When expressed with the
CC 3-ketothiolase (phbA) and acetyl-CoA-reductase (phbB) genes, a
CC poly-beta-hydroxyalkanoate (PHA), specifically poly-beta-
CC hydroxybutyrate (PHB), is expressed in the transgenic plant
CC (preferably a Brassica e.g. rape). PHB and related PHAs are
CC biodegradable thermoplastics with many useful applications.
XX
SQ Sequence 672 AA;
Query Match 24.5%; Score 34; DB 16; Length 672;
Best Local Similarity 42.9%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXLXNE 52
Db 519 WYCWLRLRHTYLQNE 532
RESULT 6
ABB05596
ID ABB05596 standard; Protein; 764 AA.
XX
AC ABB05596;
XX
DT 23-APR-2002 (first entry)
XX
DE Human testis development protein (PRTD).
XX
KW Human; testis development protein; PRTD; gene therapy.
XX
OS Homo sapiens.
XX
PN CN1318555-A.
XX
PD 24-OCT-2001.
XX
PF 11-APR-2001; 2001CN-0113501.
XX
PR 11-APR-2001; 2001CN-0113501.
XX
PA (UYNA-) UNIV NANJING MEDICAL.
XX
PI Sha J, Zhou Z, Li J;
XX
DR WPI; 2002-115092/16.
DR N-PSDB; ABA93410.
XX
PT Human testicular development relative protein gene encoded protein,

PT useful for gene therapy -

XX Claim 1; Page 1 (Claims); 7pp; Chinese.

PS

XX The present sequence represents the human testis development protein

CC designated PRTD. The PRTD gene has a cDNA sequence of 2295 base pairs

CC (bp) containing an open reading frame sequence of 764 bp from position

CC 297 to 2591, having a Genbank number of AF3111326. The present invention

CC describes: (1) utilising the PRTD gene to prepare a fusion protein;

CC (2) utilising the protein to immunise an animal and to prepare monoclonal

CC and polyclonal antibodies; and (3) utilising the PRTD gene in preparing

CC a testis development gene expressing chip. The expressed protein may be

CC useful in gene therapy for treating related diseases.

XX

SQ Sequence 764 AA;

Query Match 24.5%; Score 34; DB 23; Length 764;

Best Local Similarity 35.7%; Pred. No. 24;

Matches 5; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXXLXNE 52

DB 17 WTGWTMAQNKLFNK 30

RESULT 7

ABB70346

ID ABB70346 standard; Protein; 385 AA.

XX

AC ABB70346;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 37830.

XX

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

KW

OS Drosophila melanogaster.

XX

XX WO200171042-A2.

PN

XX

PD 27-SEP-2001.

XX

XX 23-MAR-2001; 2001WO-US09231.

PF

XX 23-MAR-2000; 2000US-191637P.

PR

PR 11-JUL-2000; 2000US-0614150.

XX

PA (PEKE) PE CORP NY.

XX

XX Venter JC, Adams M, Li PWD, Myers EW;

PI

XX WPI; 2001-656860/75.

DR

DR N-PSDB; ABL14449.

XX

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX

PS Disclosure; SEQ ID NO 37830; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 385 AA;

Query Match 23.7%; Score 33; DB 22; Length 385;

Best Local Similarity 41.2%; Pred. No. 22;

Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXXLXNEXXL 55

DB 199 WQHFYELQTTLYNEFL 215

RESULT 8

ABG29914

ID ABG29914 standard; Protein; 1213 AA.

XX

AC ABG29914;

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #29905.

XX

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder..

KW

XX Homo sapiens.

OS

XX WO200175067-A2.

PN

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US08631.

XX

XX 31-MAR-2000; 2000US-0540217.

PR

PR 23-AUG-2000; 2000US-0649167.

XX

XX (HYSE-) HYSEQ INC.

PA

XX

PI Drmanac RT, Liu C, Tang YT;

XX

XX WPI; 2001-639362/73.

DR

DR N-PSDB; AAS94101.

XX

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX

PS Claim 20; SEQ ID No 60273; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

XX Sequence 1213 AA;

Query Match 23.7%; Score 33; DB 22; Length 1213;
Best Local Similarity 35.3%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXXNEXXL 55
| | | | | | |
Db 313 QQAWHDTLTRLNREGAL 329

RESULT 9
ABB48680
ID ABB48680 standard; Protein; 1235 AA.
XX
AC ABB48680;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #1384.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides -
XX
PS Claim 6; SEQ ID No 1385; 192pp; French.
XX
CC The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1235 AA;

Query Match 23.7%; Score 33; DB 23; Length 1235;
Best Local Similarity 30.8%; Pred. No. 64;
Matches 4; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXXLN 51
| | | | |
Db 276 WSSWTHLKTSIEN 288

RESULT 10
AAO00704
ID AAO00704 standard; Protein; 136 AA.
XX
AC AAO00704;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 14596.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI80635.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 20; SEQ ID NO 14596; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 136 AA;

Query Match 22.3%; Score 31; DB 22; Length 136;
Best Local Similarity 38.5%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXXLN 51
| | | | |
Db 66 WLLWTSVQQFLRN 78

RESULT 11
AAG32774
ID AAG32774 standard; Protein; 147 AA.
XX
AC AAG32774;
XX
DT 17-OCT-2000 (first entry)
XX
DE zea mays protein fragment SEQ ID NO: 39602.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 22.3%; Score 31; DB 21; Length 147;
Best Local Similarity 45.5%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | |
Db 126 WGGWSLFQTL 136

RESULT 12
AAG32773
ID AAG32773 standard; Protein; 154 AA.
XX
AC AAG32773;
XX

DT 17-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 39601.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 22.3%; Score 31; DB 21; Length 154;
Best Local Similarity 45.5%; Pred. No. 27;

Matches 5; Conservative 0; Mismatches 0; Indels 6; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | |
Db 133 WGWWSLFQTLL 143

RESULT 13
AAG32772

ID AAG32772 standard; Protein; 167 AA.

XX

AC AAG32772;

XX

DT 17-OCT-2000 (first entry)

XX

DE Zea mays protein fragment SEQ ID NO: 39600.

XX

KW protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence; corn.

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 22.3%; Score 31; DB 21; Length 167;
Best Local Similarity 45.5%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | |
Db 146 WGWSLFQTLL 156

RESULT 14
AAU49357
ID AAU49357 standard; Protein; 399 AA.
XX
AC AAU49357;
XX
DT 13-FEB-2002 (first entry)
XX

DE Propionibacterium acnes immunogenic protein #10253.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.

XX
PF
XX
PR
PR
PR
XX
PA
XX
PI
PI
XX
DR
DR
XX
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX
SQ

20-APR-2001; 2001WO-US12865.
21-APR-2000; 2000US-199047P.
02-JUN-2000; 2000US-208841P.
07-JUL-2000; 2000US-216747P.
(CORI-) CORIXA CORP.
Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
L'maisonneuve J, Zhang Y, Jen S, Carter D;
WPI; 2001-616774/71.
N-PSDB; AAS59545.

Propionibacterium acnes polypeptides and nucleic acids useful for
vaccinating against and diagnosing infections, especially useful for
treating acne vulgaris -
Example 1; SEQ ID No 10552; 1069pp; English.

Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
polypeptides. The proteins and their associated DNA sequences are used in
the treatment, prevention and diagnosis of medical conditions caused by
P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
P. acnes is also involved in infections of bone, joints and the central
nervous system, however it is particularly involved in the inflammatory
lesions associated with acne vulgaris. A method for detecting the
presence or absence of P. acnes in a patient comprises contacting a
sample with a binding agent that binds to the proteins of the invention
and determining the amount of bound protein in the sample. The
polypeptides may be used as antigens in the production of antibodies
specific for P. acnes proteins. These antibodies can be used to
downregulate expression and activity of P. acnes polypeptides and
therefore treat P. acnes infections. The antibodies may also be used as
diagnostic agents for determining P. acnes presence, for example, by
enzyme linked immunosorbent assay (ELISA).
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

Sequence 399 AA;

Query Match 22.3%; Score 31; DB 22; Length 399;
Best Local Similarity 44.4%; Pred. No. 65;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 39 WXXWXXXQQT 47
| | | | |
Db 245 WTAWSQAQT 253

RESULT 15
AAU49903
ID AAU49903 standard; Protein; 835 AA.
XX
AC AAU49903;
XX
DT 27-FEB-2002 (first entry)
XX

DE Propionibacterium acnes immunogenic protein #10799.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.
PF
XX
XX 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
PI
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59546.
DR
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 11098; 1069pp; English.
PS
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 835 AA;
SQ
Query Match 21.9%; Score 30.5; DB 22; Length 835;
Best Local Similarity 46.7%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 39 WXXW-XXXQTXLXNE 52
| | | | |
Db 170 WANWCPNDQTVLANE 184
RESULT 16
AAB15978
ID AAB15978 standard; Protein; 99 AA.
XX
XX AAB15978;
AC
XX 05-OCT-2000 (first entry)
DT
XX E. coli proliferation associated protein sequence SEQ ID NO:335.
DE
XX Escherichia coli; E. coli; proliferation; inhibition; screening;
KW antimicrobial; bacterial growth; antisense therapy; antibacterial.
KW
XX Escherichia coli.
OS
XX WO200044906-A2.
PN
XX 03-AUG-2000.
PD
XX 27-JAN-2000; 2000WO-US02200.
PF

XX 27-JAN-1999; 99US-0117405.
PR
XX (ELIT-) ELITRA PHARM INC.
PA
XX Zyskind J, Ohlsen KL, Trawick J, Forsyth RA, Froelich JM, Carr GJ;
PI Yamamoto RT, Xu HH;
PI
XX WPI; 2000-514822/46.
DR N-PSDB; AAA65983.
DR
XX Novel polynucleotides and polypeptides associated with microorganism
PT proliferation, used to identify inhibitors of bacterial growth and
PT proliferation, for use in antisense therapy -
XX
XX Claim 11; Page 250; 316pp; English.
PS
XX AAA65809 to AAA65889 and AAA66058 to AAA66138 represent nucleotide
CC sequences derived from Escherichia coli which inhibit E. coli
CC proliferation. AAA65890 to AAA66055 and AAB15886 to AAB16040 represent
CC nucleotide and protein sequences associated with E. coli proliferation.
CC AAA66056 and AAA66057 represent primers used for sequencing E. coli
CC proliferation inhibiting nucleotide inserts in an example from the
CC present invention. Methods from the present invention can be used to
CC identify a proliferation- required gene in a microorganism, by contacting
CC a microorganism with a proliferation-required gene activity inhibitory
CC nucleic acid identified in another organism, and determining if
CC inhibition occurs in the second microorganism. The nucleic acid sequences
CC identified as being required for bacterial growth and proliferation, can
CC be used for antisense therapy for killing bacteria.
XX
XX Sequence 99 AA;
SQ
Query Match 21.6%; Score 30; DB 21; Length 99;
Best Local Similarity 28.6%; Pred. No. 30;
Matches 4; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 82 WSAWRLVKTTLKQQ 95
RESULT 17
ABB34930
ID ABB34930 standard; Peptide; 150 AA.
XX
XX ABB34930;
AC
XX 04-FEB-2002 (first entry)
DT
XX Peptide #2436 encoded by human foetal liver single exon probe.
DE
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
KW
XX Homo sapiens.
OS
XX WO200157277-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00669.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI

XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
PT
XX
PS Claim 27; SEQ ID NO 27565; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 150 AA;

Query Match 21.6%; Score 30; DB 22; Length 150;
Best Local Similarity 36.4%; Pred. NO. 44;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | : | |
Db 29 WSVWSAIKTAL 39

RESULT 18
ABB20340
ID ABB20340 standard; Protein; 150 AA.
XX
AC ABB20340;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #2339 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 15; SEQ ID No 22110; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such

CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene-expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 150 AA;

Query Match 21.6%; Score 30; DB 22; Length 150;
Best Local Similarity 36.4%; Pred. NO. 44;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | : | |
Db 29 WSVWSAIKTAL 39

RESULT 19
AAM55737
ID AAM55737 standard; Protein; 150 AA.
XX
AC AAM55737;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27842.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 27842; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 150 AA;

Query Match 21.6%; Score 30; DB 22; Length 150;

Best Local Similarity 36.4%; Pred. No. 44;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | : | |
Db 29 WSVWSAIKTAL 39

RESULT 20
ABG37643
ID ABG37643 standard; Peptide; 150 AA.
XX AC ABG37643;
XX DT 19-AUG-2002 (first entry)
XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 27308.
XX KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX OS Homo sapiens.
XX PN WO200186003-A2.
XX PD 15-NOV-2001.
XX PF 30-JAN-2001; 2001WO-US00665.
XX PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX Claim 27; SEQ ID No 27308; 634pp; English.
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method

above and (b) measuring the expression of each of the exons in several
tissues and/or cell types using hybridisation to a single exon
microarrays having a probe with the exon, where a common pattern of
expression of the exons in the tissues and/or cell types indicates that
the exons should be assigned to a single gene; a peptide comprising one
of 12011 sequences, mentioned in the specification, or encoded by the
probes/open reading frames (ORF). The probes are used for gene
expression analysis, and for identifying exons in a gene, particularly
using human lung derived mRNA and for the study of lung diseases
such as asthma, lung cancer, chronic obstructive pulmonary disease
(COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
and hyaline membrane disease. The present sequence is a peptide/protein
encoded by a single exon probe of the invention.
Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic
format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.

Sequence 150 AA;
Query Match 21.6%; Score 30; DB 23; Length 150;
Best Local Similarity 36.4%; Pred. No. 44;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | : | |
Db 29 WSVWSAIKTAL 39

RESULT 21
AAM49089
ID AAM49089 standard; Protein; 216 AA.
XX AC AAM49089;
XX DT 02-MAY-2002 (first entry)
XX DE Human flavoprotein subunit 24.
XX KW Human; flavoprotein subunit 24; recombinant production;
KW malignant tumour; cancer; blood disease; HIV infection; gene therapy;
KW human immunodeficiency virus; immune disorder; inflammatory condition;
KW cytotstatic; anti-HIV; antiinflammatory; immunomodulator.
XX OS Homo sapiens.
XX PN WO200198488-A1.
XX PD 27-DEC-2001.
XX PF 14-MAY-2001; 2001WO-CN00778.
XX PR 16-MAY-2000; 2000CN-0115731.
XX PA (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
XX PI Mao Y, Xie Y;
XX WPI; 2002-090438/12.
DR N-PSDB; ABA96872.
XX Human flavoprotein subunit 24 and encoding polynucleotide, used in
PT diagnosis and treatment of malignant tumors, hemopathy, human
PT immunodeficiency virus infection, immunological diseases and
PT inflammation -
XX Claim 1; Page 30-31; 35pp; Chinese.

CC This sequence represents human flavoprotein subunit 24. The protein
CC has a molecular weight of 24 kD. The invention relates to human
CC flavoprotein subunit 24 (AAM49089), nucleic acids encoding it (ABA96872),
CC and a method for the recombinant production of flavoprotein subunit 24.
CC The present invention additionally discloses an antagonist of
CC flavoprotein subunit 24 for therapeutic use, and an antibody which
CC specifically binds to flavoprotein subunit 24. Flavoprotein subunit 24,
CC and nucleotides which encode it may be used for treating a variety of
CC diseases, such as malignant tumours, blood diseases, HIV (human
CC immunodeficiency virus) infection, immune disorders and inflammatory
CC conditions. The protein may also be used to screen for modulators of its
CC activity or for peptide fingerprinting identification. The polynucleotide
CC can be used as a primer for nucleic acid amplification reactions or as a
CC probe for hybridisation reactions, or in producing gene chips or
CC microarrays.

XX Sequence 216 AA;

Query Match 21.6%; Score 30; DB 23; Length 216;
Best Local Similarity 45.5%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNE 52

Db 206 WSLSKTSLVNE 216

RESULT 22

AAG57493

ID AAG57493 standard; Protein; 238 AA.

XX AC AAG57493;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 74096.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149729.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 238;
Best Local Similarity 38.5%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 174 WEKWLNEQKKLAN 186

RESULT 23

AAG59921
ID AAG59921 standard; Protein; 238 AA.

XX AC AAG59921;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 77561.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 21.6%; Score 30; DB 21; Length 238;
Best Local Similarity 38.5%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXLXN 51
Db 174 WEKWLNEQKKLAN 186
RESULT 24
ABG25940
ID ABG25940 standard; Protein; 314 AA.
XX AC ABG25940;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #25931.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS90127.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID No 56299; 103pp; English.
PS The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 314 AA;
Query Match 21.6%; Score 30; DB 22; Length 314;
Best Local Similarity 45.5%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXL 49
Db 82 WINWITGOTPL 92
RESULT 25
AAG57492
ID AAG57492 standard; Protein; 331 AA.
XX AC AAG57492;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 74095.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 331;
Best Local Similarity 38.5%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXLXN 51

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 331;
Best Local Similarity 38.5%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 267 WEKWLNEQKKLAN 279

RESULT 27
AAG57491
ID AAG57491 standard; Protein; 347 AA.

XX
AC AAG57491;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 74094.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 347;
Best Local Similarity 38.5%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXTXLXN 51
| | | | |
Db 283 WEKWLNEOKKLAN 295

RESULT 28
AAG59919
ID AAG59919 standard; Protein; 347 AA.
XX
AC AAG59919;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 77559.
XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 347; *
Best Local Similarity 38.5%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXTXLXN 51
Db 283 WEKWLNEQKKLAN 295

RESULT 29
AAU24598
ID AAU24598 standard; Protein; 349 AA.
XX AAU24598;

XX 18-DEC-2001 (first entry)
XX Human olfactory receptor AOLF89.
DE Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW food additive; cosmetic; fragrance; pharmaceutical additive.
XX Homo sapiens.
OS WO200168805-A2.
PN
XX

PD 20-SEP-2001.
XX
PF 13-MAR-2001; 2001WO-US07771.
XX
PR 13-MAR-2000; 2000US-0188914.
PR 24-MAR-2000; 2000US-0192033.
PR 12-APR-2000; 2000US-0198474.
PR 24-APR-2000; 2000US-0199335.
PR 26-MAY-2000; 2000US-0207702.
PR 23-JUN-2000; 2000US-0213849.
PR 16-AUG-2000; 2000US-0226534.
PR 07-SEP-2000; 2000US-0230732.
PR 07-FEB-2001; 2001US-0266862.
XX
PA (SENO-) SENOMYX INC.
XX
PI Zozulya S;
XX
DR WPI; 2001-570867/64.
DR N-PSDB; AAS42291.
XX
PT Nucleic acids encoding human olfactory G protein-coupled receptors,
PT useful for screening for compounds involved in olfactory sensation,
PT where the compounds can be used in the food, pharmaceutical and
PT cosmetic industries to customise odours -
XX
PS Claim 60; Page 120; 319pp; English.
XX
CC The invention relates to nucleic acids encoding human olfactory
CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC specifically recognise molecules, odourants, that elicit specific
CC olfactory sensation. The human olfactory receptors and polynucleotides
CC encoding them are useful for screening a library of chemical compounds
CC for compounds that are involved in olfactory sensation. Modulators of
CC their activity are useful for pharmacological and genetic modulation of
CC olfactory signalling pathways. Therefore, they can be used in the food,
CC pharmaceutical and cosmetic industries to customise odours and
CC fragrances. The present sequence is a human olfactory receptor of the
CC invention.
XX
SQ Sequence 349 AA;

Query Match 21.6%; Score 30; DB 22; Length 349;
Best Local Similarity 42.9%; Pred. No. 97;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNEXXL 55
Db 27 WQKNQTSLADEFIL 40

RESULT 30
AAG06011
ID AAG06011 standard; Protein; 363 AA.
XX

AC AAG06011;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2628.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
PN
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
PF
XX

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 20-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 363;
Best Local Similarity 36.4%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
Db 186 WAAWFLIQTKM 196

RESULT 31
AAG47673
ID AAG47673 standard; Protein; 363 AA.
XX AC AAG47673;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 60112.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 363;

Best Local Similarity 36.4%; Pred. No. 1e+02; Mismatches 1; Indels 0; Gaps 0;

Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXQTXL 49
| | | | :
Db 186 WAAWFIQTKM 196

RESULT 32

AAG06010
ID AAG06010 standard; Protein; 369 AA.

XX AC AAG06010;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 2627.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.

PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 369;
Best Local Similarity 36.4%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | | | |
Db 192 WAAWFIIQTKM 202

RESULT 33
AAG47672
ID AAG47672 standard; Protein; 369 AA.

XX AC AAG47672;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 60111.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 369;
Best Local Similarity 36.4%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 39 WXXWXXXQTXL 49
| | | | |
Db 192 WAAWFIQTKM 202

RESULT 34

AAG06009
ID AAG06009 standard; Protein; 374 AA.

XX AC AAG06009;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 2626.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR	26-OCT-1999;	99US-0161360.	PR	01-JUN-1999;	99US-0137222.
PR	26-OCT-1999;	99US-0161361.	PR	03-JUN-1999;	99US-0137528.
PR	28-OCT-1999;	99US-0161920.	PR	04-JUN-1999;	99US-0137502.
PR	28-OCT-1999;	99US-0161992.	PR	07-JUN-1999;	99US-0137724.
PR	28-OCT-1999;	99US-0161993.	PR	08-JUN-1999;	99US-0138094.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;	99US-0138540.
Query Match 21.6%; Score 30; DB 21; Length 374;					
Best Local Similarity 36.4%; Pred. NO. 1e+02;					
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;					
QY	39 WXXWXXXQTXL 49				
	:				
Db	197 WAAWFIIQTKM 207				
RESULT 35					
AAG47671					
ID	AAG47671 standard; Protein; 374 AA.				
XX					
AC	AAG47671;				
XX					
DT	18-OCT-2000 (first entry)				
XX					
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 60110.				
XX					
KW	protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
XX					
OS	Arabidopsis thaliana.				
XX					
PN	EP1033405-A2.				
XX					
PD	06-SEP-2000.				
XX					
PF	25-FEB-2000; 2000EP-0301439.				
XX					
PR	25-FEB-1999;	99US-0121825.	PR	13-JUL-1999;	99US-0143542.
PR	05-MAR-1999;	99US-0123180.	PR	14-JUL-1999;	99US-0143624.
PR	09-MAR-1999;	99US-0123548.	PR	15-JUL-1999;	99US-0144005.
PR	23-MAR-1999;	99US-0125788.	PR	16-JUL-1999;	99US-0144085.
PR	25-MAR-1999;	99US-0126264.	PR	16-JUL-1999;	99US-0144086.
PR	29-MAR-1999;	99US-0126785.	PR	19-JUL-1999;	99US-0144325.
PR	01-APR-1999;	99US-0127462.	PR	19-JUL-1999;	99US-0144331.
PR	06-APR-1999;	99US-0128234.	PR	19-JUL-1999;	99US-0144332.
PR	08-APR-1999;	99US-0128714.	PR	19-JUL-1999;	99US-0144333.
PR	16-APR-1999;	99US-0129845.	PR	19-JUL-1999;	99US-0144334.
PR	19-APR-1999;	99US-0130077.	PR	19-JUL-1999;	99US-0144335.
PR	21-APR-1999;	99US-0130449.	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	99US-0130510.	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	99US-0130891.	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-01511303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 374;
Best Local Similarity 36.4%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 39 WXXWXXXTXL 49
Db 197 WAAWFIIQTKM 207
RESULT 36
AAG22294
ID AAG22294 standard; Protein; 382 AA.
XX AC AAG22294;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 25166.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 382;
Best Local Similarity 28.6%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXQTXLXNE 52
| | | | |
Db 81 WSEWAEQNSLSED 94

RESULT 37
AAG22293
ID AAG22293 standard; Protein; 386 AA.
XX
AC AAG22293;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 25165.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 386;
Best Local Similarity . 28.6%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | | | | | | |
Db 85 WSEWEAEQNSLSED 98

RESULT 38
AAG22292
ID AAG22292 standard; Protein; 405 AA.
XX
AC AAG22292;

XX 17-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 25164.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 405;
Best Local Similarity 28.6%; Pred. No. 1.1e+02;
Matches 4; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXQTXLXNE 52
| | | | |
Db 104 WSEWEAEQNSLSED 117

RESULT 39
AAG38504
ID AAG38504 standard; Protein; 526 AA.
XX
AC AAG38504;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47511.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW	termination sequence.		
XX		PR	28-JUN-1999;
OS	Arabidopsis thaliana.	PR	29-JUN-1999;
XX		PR	30-JUN-1999;
PN	EP1033405-A2.	PR	01-JUL-1999;
XX		PR	01-JUL-1999;
XX		PR	02-JUL-1999;
PD	06-SEP-2000.	PR	06-JUL-1999;
XX		PR	08-JUL-1999;
PF	25-FEB-2000; 2000EP-0301439.	PR	09-JUL-1999;
XX		PR	12-JUL-1999;
PR	25-FEB-1999; 99US-0121825.	PR	13-JUL-1999;
PR	05-MAR-1999; 99US-0123180.	PR	14-JUL-1999;
PR	09-MAR-1999; 99US-0123548.	PR	15-JUL-1999;
PR	23-MAR-1999; 99US-0125788.	PR	16-JUL-1999;
PR	25-MAR-1999; 99US-0126264.	PR	16-JUL-1999;
PR	29-MAR-1999; 99US-0126785.	PR	19-JUL-1999;
PR	01-APR-1999; 99US-0127462.	PR	19-JUL-1999;
PR	06-APR-1999; 99US-0128234.	PR	19-JUL-1999;
PR	08-APR-1999; 99US-0128714.	PR	19-JUL-1999;
PR	16-APR-1999; 99US-0129845.	PR	19-JUL-1999;
PR	19-APR-1999; 99US-0130077.	PR	20-JUL-1999;
PR	21-APR-1999; 99US-0130449.	PR	20-JUL-1999;
PR	23-APR-1999; 99US-0130510.	PR	20-JUL-1999;
PR	23-APR-1999; 99US-0130891.	PR	21-JUL-1999;
PR	28-APR-1999; 99US-0131149.	PR	21-JUL-1999;
PR	30-APR-1999; 99US-0132048.	PR	21-JUL-1999;
PR	30-APR-1999; 99US-0132407.	PR	22-JUL-1999;
PR	04-MAY-1999; 99US-0132484.	PR	22-JUL-1999;
PR	05-MAY-1999; 99US-0132485.	PR	22-JUL-1999;
PR	06-MAY-1999; 99US-0132486.	PR	22-JUL-1999;
PR	06-MAY-1999; 99US-0132487.	PR	22-JUL-1999;
PR	07-MAY-1999; 99US-0132863.	PR	23-JUL-1999;
PR	11-MAY-1999; 99US-0134256.	PR	23-JUL-1999;
PR	14-MAY-1999; 99US-0134218.	PR	23-JUL-1999;
PR	14-MAY-1999; 99US-0134219.	PR	26-JUL-1999;
PR	14-MAY-1999; 99US-0134221.	PR	27-JUL-1999;
PR	14-MAY-1999; 99US-0134370.	PR	27-JUL-1999;
PR	18-MAY-1999; 99US-0134768.	PR	27-JUL-1999;
PR	19-MAY-1999; 99US-0134941.	PR	28-JUL-1999;
PR	20-MAY-1999; 99US-0135124.	PR	27-JUL-1999;
PR	21-MAY-1999; 99US-0135353.	PR	02-AUG-1999;
PR	24-MAY-1999; 99US-0135629.	PR	02-AUG-1999;
PR	25-MAY-1999; 99US-0136021.	PR	02-AUG-1999;
PR	27-MAY-1999; 99US-0136392.	PR	03-AUG-1999;
PR	28-MAY-1999; 99US-0136782.	PR	04-AUG-1999;
PR	01-JUN-1999; 99US-0137222.	PR	04-AUG-1999;
PR	03-JUN-1999; 99US-0137528.	PR	05-AUG-1999;
PR	04-JUN-1999; 99US-0137502.	PR	05-AUG-1999;
PR	07-JUN-1999; 99US-0137724.	PR	06-AUG-1999;
PR	08-JUN-1999; 99US-0138094.	PR	06-AUG-1999;
PR	10-JUN-1999; 99US-0138540.	PR	09-AUG-1999;
PR	10-JUN-1999; 99US-0138847.	PR	09-AUG-1999;
PR	14-JUN-1999; 99US-0139119.	PR	10-AUG-1999;
PR	16-JUN-1999; 99US-0139452.	PR	11-AUG-1999;
PR	16-JUN-1999; 99US-0139453.	PR	12-AUG-1999;
PR	17-JUN-1999; 99US-0139492.	PR	13-AUG-1999;
PR	18-JUN-1999; 99US-0139454.	PR	13-AUG-1999;
PR	18-JUN-1999; 99US-0139455.	PR	16-AUG-1999;
PR	18-JUN-1999; 99US-0139456.	PR	16-AUG-1999;
PR	18-JUN-1999; 99US-0139457.	PR	17-AUG-1999;
PR	18-JUN-1999; 99US-0139458.	PR	18-AUG-1999;
PR	18-JUN-1999; 99US-0139459.	PR	20-AUG-1999;
PR	18-JUN-1999; 99US-0139460.	PR	20-AUG-1999;
PR	18-JUN-1999; 99US-0139461.	PR	23-AUG-1999;
PR	18-JUN-1999; 99US-0139462.	PR	23-AUG-1999;
PR	18-JUN-1999; 99US-0139463.	PR	25-AUG-1999;
PR	18-JUN-1999; 99US-0139750.	PR	26-AUG-1999;
PR	18-JUN-1999; 99US-0139763.	PR	27-AUG-1999;
PR	21-JUN-1999; 99US-0139817.	PR	27-AUG-1999;
PR	22-JUN-1999; 99US-0139899.	PR	27-AUG-1999;
PR	23-JUN-1999; 99US-0140353.	PR	30-AUG-1999;
PR	23-JUN-1999; 99US-0140354.	PR	31-AUG-1999;
PR	24-JUN-1999; 99US-0140695.	PR	01-SEP-1999;
		PR	07-SEP-1999;
			99US-0140823.
			99US-0140991.
			99US-0141287.
			99US-0141842.
			99US-0142154.
			99US-0142055.
			99US-0142390.
			99US-0142803.
			99US-0142920.
			99US-0142977.
			99US-0143542.
			99US-0143624.
			99US-0144005.
			99US-0144085.
			99US-0144086.
			99US-0144325.
			99US-0144331.
			99US-0144332.
			99US-0144333.
			99US-0144334.
			99US-0144335.
			99US-0144352.
			99US-0144632.
			99US-0144884.
			99US-0144814.
			99US-0145086.
			99US-0145088.
			99US-0145085.
			99US-0145087.
			99US-0145089.
			99US-0145192.
			99US-0145145.
			99US-0145218.
			99US-0145224.
			99US-0145276.
			99US-0145913.
			99US-0145918.
			99US-0145919.
			99US-0145951.
			99US-0146386.
			99US-0146388.
			99US-0146389.
			99US-0147038.
			99US-0147204.
			99US-0147302.
			99US-0147192.
			99US-0147260.
			99US-0147303.
			99US-0147416.
			99US-0147493.
			99US-0147935.
			99US-0148171.
			99US-0148319.
			99US-0148341.
			99US-0148565.
			99US-0148684.
			99US-0149368.
			99US-0149175.
			99US-0149426.
			99US-0149722.
			99US-0149723.
			99US-0149929.
			99US-0149902.
			99US-0149930.
			99US-0150566.
			99US-0150884.
			99US-0151065.
			99US-0151066.
			99US-0151080.
			99US-0151303.
			99US-0151438.
			99US-0151930.
			99US-0152363.

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 526;
Best Local Similarity 45.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 42 WXXXQTXLXNE 52
I I I I I I
Db 81 WRVKQTDIANE 91

RESULT 40
AAG38503
ID AAG38503 standard; Protein; 536 AA.

XX AC AAG38503;

DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 47510.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.

PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 536;
Best Local Similarity 45.5%;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 42 WXXXQTXLXNE 52
| | | : | |
Db 91 WRVKQTDIANE 101

RESULT 41
AAB93159
ID AAB93159 standard; Protein; 882 AA.
XX
AC AAB93159;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:12081.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX

PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 12081; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 882 AA;

Query Match 21.6%; Score 30; DB 22; Length 882;
Best Local Similarity 36.4%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXL 49
| | : | |
Db 785 WNAWTALEYL 795

RESULT 42
AAG46753
ID AAG46753 standard; Protein; 1730 AA.
XX
AC AAG46753;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58853.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 1730;
Best Local Similarity 28.6%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | | | | | | |
Db 71 WSEWEAEQNSLSED 84

RESULT 43

AAG46752
ID AAG46752 standard; Protein; 1734 AA.

XX
AC AAG46752;

XX
DT 18-OCT-2000 (first entry)

XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58852.

XX
KW protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX
OS Arabidopsis thaliana.

XX
PN EP1033405-A2.

XX
PD 06-SEP-2000.

XX
PF 25-FEB-2000; 2000EP-0301439.

XX
PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	23-APR-1999;	99US-0130891.
PR	28-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	06-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-01511066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-01511303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.

PR 14-OCT-1999; 99US-01593330.
PR 14-OCT-1999; 99US-01593331.
PR 14-OCT-1999; 99US-01596337.
PR 14-OCT-1999; 99US-01596338.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.6%; Score 30; DB 21; Length 1734;
Best Local Similarity 28.6%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXNE 52
| | | | |
Db 75 WSEWAEQNSLSED 88

RESULT 44
ABB32215
ID ABB32215 standard; Peptide; 58 AA.

XX
AC ABB32215;

DT 01-FEB-2002 (first entry)

XX Peptide #4866 encoded by breast cell single exon nucleic acid probe.
DE Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer.
KW Homo sapiens.

OS
XX
PN WO200157271-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00662.

PF 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-496933/54.

XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX

PS Claim 27; SEQ ID NO 15183; 327pp + sequence listing; English.

XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 58 AA;

Query Match 20.9%; Score 29; DB 22; Length 58;

Best Local Similarity 30.8%; Pred. No. 31;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 19 WTTWEPAATLIFN 31

RESULT 45
ABB37467
ID ABB37467 standard; Peptide; 58 AA.

XX
AC ABB37467;

DT 04-FEB-2002 (first entry)

XX Peptide #4973 encoded by human foetal liver single exon probe.

DE Human; foetal liver; gene expression; single exon nucleic acid probe.

KW Homo sapiens.

OS WO200157277-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00669.

PF 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -

XX Claim 27; SEQ ID NO 30102; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 58 AA;

Query Match 20.9%; Score 29; DB 22; Length 58;
Best Local Similarity 30.8%; Pred. No. 31;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | : |
Db 19 WTTWEPAATLIFN 31

RESULT 46
ABB22761
ID ABB22761 standard; Protein; 58 AA.
XX
AC ABB22761;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #4760 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 15; SEQ ID No 24531; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 58 AA;

Query Match 20.9%; Score 29; DB 22; Length 58;
Best Local Similarity 30.8%; Pred. No. 31;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | : |
Db 19 WTTWEPAATLIFN 31

RESULT 47
AAM58139
ID AAM58139 standard; Protein; 58 AA.
XX
AC AAM58139;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30244
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 30244; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, of
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 58 AA;

Query Match 20.9%; Score 29; DB 22; Length 58;
Best Local Similarity 30.8%; Pred. No. 31;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | : |
Db 19 WTTWEPAATLIFN 31

RESULT 48

AAM70588
ID AAM70588 standard; Protein; 58 AA.
XX
AC AAM70588;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 30894.
XX
DE Human bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
KW
XX Homo sapiens.
OS
XX
XX WO200157276-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00668.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
PR 26-MAY-2000; 2000US-0207456.
PR
PR 30-JUN-2000; 2000US-0608408.
PR
PR 03-AUG-2000; 2000US-0632366.
PR
PR 21-SEP-2000; 2000US-0234687.
PR
PR 27-SEP-2000; 2000US-0236359.
PR
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-488900/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
PT
XX
XX Example 4; SEQ ID NO: 30894; 658pp + Sequence Listing; English.
PS
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
XX
SQ Sequence 58 AA;
Query Match 20.9%; Score 29; DB 22; Length 58;
Best Local Similarity 30.8%; Pred. No. 31;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXLXN 51
| | | : |
Db 19 WTTWEPAATLIFN 31
RESULT 49
AAM18425
ID AAM18425 standard; Protein; 58 AA.
XX
AC AAM18425;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #4859 encoded by probe for measuring cervical gene expression.
XX Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
KW
XX Homo sapiens.
OS
XX
XX WO200157278-A2.
PN

XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00670.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
PR 26-MAY-2000; 2000US-0207456.
PR
PR 30-JUN-2000; 2000US-0608408.
PR
PR 03-AUG-2000; 2000US-0632366.
PR
PR 21-SEP-2000; 2000US-0234687.
PR
PR 27-SEP-2000; 2000US-0236359.
PR
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-488901/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
PT
XX
XX Claim 27; SEQ ID No 23251; 487pp; English.
PS
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 58 AA;
Query Match 20.9%; Score 29; DB 22; Length 58;
Best Local Similarity 30.8%; Pred. No. 31;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 39 WXXWXXXQTXLXN 51
| | | : |
Db 19 WTTWEPAATLIFN 31
RESULT 50
AAM30901
ID AAM30901 standard; Protein; 58 AA.
XX
AC AAM30901;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #4938 encoded by probe for measuring placental gene expression.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
KW
XX Homo sapiens.
OS
XX
XX WO200157272-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00663.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
PR 26-MAY-2000; 2000US-0207456.
PR
PR 30-JUN-2000; 2000US-0608408.
PR
PR 03-AUG-2000; 2000US-0632366.
PR
PR 21-SEP-2000; 2000US-0234687.
PR

PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
PT
XX
PS Claim 27; SEQ ID No 31170; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 58 AA;

Query Match 20.9%; Score 29; DB 22; Length 58;
Best Local Similarity 30.8%; Pred. No. 31;
Matches 4; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 39 WXXWXXXQTXLXN 51
| | | | |
Db 19 WTTWEPAAATLIFN 31

Search completed: July 18, 2003, 15:57:28
Job time : 71 secs